

STIC-Biotech/ChemLib

From: Helmer, Georgia
Sent: Wednesday, August 28, 2002 10:51 AM
To: STIC-Biotech/ChemLib
Subject: FW: seq search 09/810861

-----Original Message-----

From: Helmer, Georgia
Sent: Wednesday, August 28, 2002 10:33 AM
T : STIC-ILL
Subject: seq search 09/810861

Antigen

Please do DNA sequence searches, and oligo searches, for SEQ ID NOs 1 thru 5 of this case.

Thanks in advance!

Georgia L. Helmer Ph.D.
Patent Examiner
Crystal Mall 1, 9D14
AU 1638
703-308-7023
Georgia.Helmer@USPTO.gov
mailbox 9e12

12/1/02

POINT OF CONTACT:
PAUL SCHULWITZ
TECHNICAL INFO. SPECIALIST
CM1 6B06 TEL. (703) 305-1954

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 8/30
Date Completed: 9/3/02
Searcher Prep/Review: 10
Clerical: _____
Online time: 10

TYPE OF SEARCH:
NA Sequences: 4
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:41:47 ; Search time 9324.54 Seconds
(without alignments)
3871.320 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atgagggcccgagtgctc.....aatgtcagcgccacctga 1725

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1725	100.0	1725	6	AX275256	AX275256 Sequence
2	1722	99.8	2218	6	HUMACHE	M55040 Human acety
3	1722	99.8	2256	6	AR070205	AR070205 Sequence
4	1722	99.8	3016	6	AR070207	AR070207 Sequence
5	1722	99.8	3096	6	AR070206	AR070206 Sequence
6	1712	99.2	5767	6	AX275254	AX275254 Sequence
7	1712	99.2	14446	6	AX275255	AX275255 Sequence
8	1070	62.0	1355	9	HUMACHE01	L22559 Human acety
9	1070	62.0	4185	9	HUMACHEB	L42812 Homo sapien
10	1070	62.0	172358	9	AC011895	AC011895 Homo sapi
11	1070	62.0	194681	2	AC084057	AC084057 Homo sapi
12	781	45.3	34921	9	HSARF002993	AF002993 Homo sapi
13	489	28.3	779	9	HUMACHE02	L22560 Human acety
14	170	9.9	510	9	HUMACHE03	L22561 Human acety
15	164	9.5	117962	9	AF312032	AF312032 Homo sapi
16	60	3.5	3958	4	AF053485	AF053485 Felis cat
17	53	3.1	1947	4	OCU05036	U05036 Oryctolagus
18	48	2.8	503	4	BTACHE2	AF061814 Bos tauru
19	47	2.7	984	4	BTACHE1	AF061813 Bos tauru
20	33	1.9	2089	6	AX306139	AX306139 Sequence
21	33	1.9	2089	10	MMACHE	X56518 Mouse mRNA
22	33	1.9	296820	10	AF312033	AF312033 Mus muscu
23	32	1.9	2066	10	S50879	S50879 acetylcho11
24	30	1.7	535	9	HUMACHEA	M76539 Human acety
25	28	1.6	183865	9	AC009084	AC009084 Homo sapi
26	28	1.6	206102	9	AC027131	AC027131 Homo sapi
27	27	1.6	1538	6	AX180336	AX180336 Sequence
28	25	1.4	2523	6	AX179306	AX179306 Sequence
29	25	1.4	3993	10	RNU41662	U41662 Rattus norv
30	30	1.4	65064	2	AC098752	AC098752 Rattus no
31	25	1.4	146756	2	AC002348	AC002348 Homo sapi
32	25	1.4	185010	2	AC022514	AC022514 Mus muscu
33	25	1.4	190358	2	AC016876	AC016876 Homo sapi
34	25	1.4	222871	2	AC108094	AC108094 Homo sapi
35	25	1.4	234182	10	AL603707	AL603707 Mouse DNA
36	24	1.4	840	10	MUSACHEA	M76540 Mouse acety
37	23	1.3	667	10	AF151982	AF151982 Rattus no
38	23	1.3	182844	2	AL671857	AL671857 Mus muscu
39	22	1.3	3945	10	AB043123	AB043123 Mus muscu
40	22	1.3	4011	10	AB043125	AB043125 Mus muscu
41	22	1.3	4014	10	AB043124	AB043124 Mus muscu
42	22	1.3	4173	10	AB005662	AB005662 Mus muscu
43	22	1.3	4545	10	AF262046	AF262046 Mus muscu
44	22	1.3	5443	10	AF178636	AF178636 Mus muscu
45	22	1.3	5532	10	AF178637	AF178637 Mus muscu

ALIGNMENTS

RESULT	1	LOCUS	AX275256	Sequence	5 from Patent WO0171014.	DNA	1725 bp	linear	PAT	29-OCT-2001
--------	---	-------	----------	----------	--------------------------	-----	---------	--------	-----	-------------

DEFINITION AX275256
ACCESSION AX275256
VERSION AX275256.1
KEYWORDS GI:16547676
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PATENT: WO 0171014-A 5 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; MOR, Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason, Hugh S. (US)

FEATURES
source
Location/Qualifiers
1..1725

/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic human acetylcholinesterase gene optimized"

for expression in plants"
BASE COUNT 273 a 553 c 571 g 328 t
ORIGIN

Query Match 100.0%; Score 1725; DB 6; Length 1725;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1725; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagagcccccagctgtctctgtcacacagccctccctgtgtcccccactctctctc 60
Db 1 ATGAGGCCCCCGACAGTCTGTGCACACAGCCCTTCCCTGGCTTCCCACTCTCTCCTC 60
QY 61 ctcctctgtctcctgtgtgagagagtggtgggtgagggccgggagagatgcagagctgtg 120
Db 61 CTCTCTGTGCTCTGGGTGGAGAGAGTGGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 120
QY 121 gtgacggtgcgtgtgggagccggtctgcggggcattccgctgaagacccccggggccctgtc 180
Db 121 GTGACGGGTGCGTGGGGGCGGCTGCGGGGCATTCGCCTGAAGACCCCCGGGGCCCTGTGTC 180
QY 181 tctgtcttctctgtgcatccccccttgcggaagcccaatggaaccccgctcgtcttctgcca 240
Db 181 TCTGCTTCTCTGGGCAATCCCCCTTTGCGGAGCCACCATGGGACCCCGTGGCTTCTGCA 240
QY 241 ccggaagcccaagcagccttgctcagagggtgtgtagacgtctacaaccttcagagtgctgc 300
Db 241 CCGGAGCCCAAGCAGCCTTGGTTCAGGGGTGTAGACGCTTACAACCTTCCAGAGTGTCTGC 300
QY 301 taccaatatgtgacacccctatacccaagttttgagggcaccgagatgttgaaaccccaac 360
Db 301 TACCAATATGTGACACCCCTATACCCAGGTTTGAAGGGACCCGAGATGTGAACCCCAAC 360
QY 361 cgtgagctgagcgagactgctgttacctcaacggtgtgacaccataccccgggcttaca 420
Db 361 CGTAGAGTGAAGGAGACTGCTGTACTCTCAACGCTGTGAGACACCATACCCCGGCTTACA 420
QY 421 tcccccaaccctgtctctcgtctgtgatactatgtgggtgtgcttctacagtgtgggctctcc 480
Db 421 TCCCCCACCCCTGTCTCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTCC 480
QY 481 ttggaagctgaagatggccgcttctgtgtacagccggaagagactgtgtgtgtccatg 540
Db 481 TTGGAAGCTGTACGATGGCCGCTTCTGTGTACAGGCCGAGAGACTGTGTGTCTCATG 540
QY 541 aactaacggtgtggagccttlttgctctgtgcccctgcgggagccgagagccccgggc 600
Db 541 AACTAACGGGTGGAGCCTTTGGCTTCTGTGCCCTTGCCGGGAGCCGAGAGGCCCGGGGC 600
QY 601 aatgtgggtctcctgtgatacagaggtgagccctgcagtggtgcagagaaagtggcagcc 660
Db 601 AATGTGGGTCTCCTGGATCAAGAGCTGGCCCTGCAGTGGGTGCAGAGAAACGTGGCAGCC 660
QY 661 ttcgggggtgacccgacatcagtgacgctgtttggygagagcgcggyagccgctcgtgtg 720
Db 661 TTCGGGGGTGACCCGACATCAGTACGCTGTGGGGAGAGAGCGGGAGCCGCTCGGTG 720
QY 721 ggcatgcaacctgtgtcccccggccagccggggcctgttccacaagggccgtgtcagagc 780
Db 721 GGCATGCACCTGTGTCTCCCGCCAGCCGGGGCTGTTCACACAGGGCCGTGTGCAGAGC 780
QY 781 ggtgcccccaatggaacctgtggcaggtgtggcatgtagagagccccgtgcagggccagc 840
Db 781 GGTGCCCCCAATGGAACCTTGGGCCACGGTGGCATGGGAGAGGCCCTGCCAGGGCCACG 840
QY 841 cagctggcccaacctgtgggtgtgtctcctcagggcggaactgtgtggaatgacacagagctg 900
Db 841 CAGCTGGCCCAACCTGTGGGTGTCTCTCCAGGGCGCACTGTGGGAATGACACAGAGCTG 900
QY 901 gttagcctgtccttggaacagcagcagcgcaagtcctgtgtgaaccaagaaatgcagtgctg 960
Db 901 GTAGCCTGTCTGGACACGACGACGCGCAGGTCCTGTGTGAACCAACGAATGGCACGTCGTG 960

QY 961 cctcaagaagcgtcttccgggttctcctctcgtgtgcctgtgtagatgagacttccctcagt 1020
Db 961 CCTCAAGAAAGCGTCTCCGGTCTCTCCTGTGCTGTGTAGATGAGACTTCTCTCAGT 1020
QY 1021 gacacccccagagggccctcatcaacgcgggagagacttccacgacctgcagtgctgtgtgt 1080
Db 1021 GACACCCCCAGAGGCCCTCATCAACGCGGAGACTTCCACGGCCTGCAGGTGCTGTGGGT 1080
QY 1081 gtgtgtgaagatgagggctcgtatttctgtgttaagggggccccagagcttcaagcaaaagc 1140
Db 1081 GTGTGAAGATGAGGGCTCGTATTCTGTGTTAAGGGGGCCCCAGAGCTTCAAGCAAAAGAC 1140
QY 1141 aacaggtctctcatcagccgggcccaggttctctggccggggtgcgggtcggggttccccag 1200
Db 1141 AACAGTCTCTCATCAGCCGGCCGAGTTCCTGGCCGGGTGCGGGTCTGCCAG 1200
QY 1201 gtaagtgaacctgagccgaggtgtgtctctgcatcaacagactgtgtcatcccgag 1260
Db 1201 GTAAGTGACCTGGCAGCCGAGGCTGTGTCTCTGCATTAACAGACTGGCTGCATCCCGAG 1260
QY 1261 gaccggcagcgcctgagggagggccctgagcgaatgtgtgtggcgacacacaatgtctgtgc 1320
Db 1261 GACCCGGCAGCCCTGAGGGAGGCCCCCTGAGCGATGTGTGGCGCAGCACAAATGTCTGTGC 1320
QY 1321 cccgtggccagctgtgctggcgactgtgctgcccaagggtgccgggtctacgctctacgtc 1380
Db 1321 CCGGTGGCCAGCTGTGCTGGCGGACTGTGCTGCCAGGGGTGCCGGGTCTACGCTTACGTC 1380
QY 1381 ttggaacacgtgtcttccacgctctcctgtgcccctgtgagtgggtgtccccacggtctac 1440
Db 1381 TTTGAACACCGTGTCTTCCACGCTCTCTGGCCCTGTGTGATGGGGTGTGCCCAAGGCTTAC 1440
QY 1441 gagatcgaagttcatcttctggatccccctgtgacccctctcgaactacagacagcagagag 1500
Db 1441 GAGATCGAGTTCACTTTGGGATCCCCCTGAGCCCTCTCGAAACTACAGCGCAGAGAG 1500
QY 1501 aaaatcttcgcccagcgactgtatgcatactgtggccaacttgcgccgcagagggatccc 1560
Db 1501 AAAATCTTGCCCAGGACTGATGCATCTGGGCCAACTTTGCCCGCACAGGGGATCCC 1560
QY 1561 aatgagccccgagagcccccaagggccccacaatgtgccccgtlacagcgcggtcagcag 1620
Db 1561 AATGAGCCCCGAGAGCCCAAGGCCCCACAATGGCCCCCTATACAGCGCGGGGCTCAGCAG 1620
QY 1621 taagttagctgtgacctgtgcgcgtgtgaggtgtgcggtgggtgcgcgcccagggcctgc 1680
Db 1621 TACGTTAGCTGTGACCTGCGGCCCTGTGAGAGTGGCGGGGGCTGTGCGGCCCAAGGCTGC 1680
QY 1681 gcctctggaaccgcttccctccccaatgtctcagcgccacctga 1725
Db 1681 GCCTTCTGAACCGCTTCTCTCCCAAAATGTCTCAGCGCCACCTGA 1725

RESULT 2
HUMACHE HUMACHE 2218 bp mRNA linear PRI 09-JUN-1995
LOCUS
DEFINITION Human acetylcholinesterase (ACHE) mRNA, complete cds.
ACCESSION M55040
VERSION M55040.1 GI:177974
KEYWORDS acetylcholinesterase.
SOURCE Human 21-week old fetus DNA, and cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2218)
AUTHORS Soreq,H.E., Ben-Aziz,R., Prody,C.A., Seidman,S., Gnatt,A.,
Neville,L., Lieberman-Hurwitz,J., Lev-Lehman,E., Ginzberg,D.,
Lapidot-Lifson,Y. and Zakut,H.
TITLE Molecular cloning and construction of the coding region for human
acetylcholinesterase reveals a G + C-rich attenuating structure
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9688-9692 (1990)
MEDLINE 91088577
FEATURES Location/Qualifiers

source	1. . 2218
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/map="7q22"
	/dev_stage="21-week-old fetus"
	/tissue_1lb="lambda-glt10"
mRNA	1. . 2218
	/gene="ACHE"
	/note="G00-118-746"
	/product="acetylcholinesterase"
gene	1. . 2218
	/gene="ACHE"
	157. . 2001
CDS	/gene="ACHE"
	/EC_number="3.1.1.7"
	/codon_start=1
	/product="acetylcholinesterase"
	/protein_id="AA68151.1"
	/db_xref="GI:177975"
	/db_xref="GDB:G00-118-746"
	/translation="MRPQCLHTPSLASPLLILLMLLGGVGAGREDALLTVR GRLRGIRLKTTPGGPVSAFLGPFAEPPMGPRRFLPEPKQPWGSVDAITTFQSVCYQ YVDLYPGFECTEMNPRELSEDCLYINWTYP RP TSPVLWTYGGFYSGASS LDVYDGRFLVQAERTVLVSMTNRYGAFFGLALPGSREAPGNVGLDORIALQWOENV AAFSGDPTSVTLFGESAGAASVGMHLSPSRGLFHRAYLOSAPNGMATVMGEAR RRATQLAHLVGCPPGGTGNDTELVACLRPAQVLNVHEWHVLPQESVERFSVPV DGDFLSDTPEALLINAGDEHGLQVLGVVKDEGSFYLVYGAPEGFSKNESSLISRAEFLLA GVRGVPQVSDLAEAVALHYTDWLHPEDPARLREALSDVDGHNYCPVAOLAGRLLA AGCARVAVVFEEHRASTLSWPIIMMGVPHGYEIEIFIGIPIDPSRNTAEKIFAQRLLM RYWANFARTGDPNRPDKAPQMPPTAGAOOVSIDLRLPLEVRGRLAQACAFWNRF LPKL SATDTLDEAERQWKAEFHMSYMWVNKNQFDHYSKDRCSDL" 229. . 2291
s1g_peptide	/gene="ACHE"
	/note="G00-118-746; does not fit consensus"
mat_peptide	310. . 1998
	/gene="ACHE"
	/EC_number="3.1.1.7"
	/note="G00-118-746"
	/product="acetylcholinesterase"
BASE COUNT	355 a 755 c 680 g 428 t
ORIGIN	
Query Match	99.8%; Score 1722; DB 9; Length 2218;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1722; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 atgagcccccgagctgtcctgcacacacgacctccctggcttcccactctctc 60
Db	157 ATGAGGCCCGCAGCTGTCTGCACACGCGCTCCCTGCGCTTCCCACACTCTCTCTC 216
Qy	61 ctccctgcctcctgtgtgaggagtgtgggcttgagggccgggagatgcagagctgtc 120
Db	217 CTCCTCTGGCTCCTGGGTGAGAGAGTGGGGGCTGAGGGCGGGAAGATGCAGAGCTGCTG 276
Qy	121 gtgacggtgcgttggggccggtgcggygcattgcctgaagaacccccggggccctgtc 180
Db	277 GTGACGGGTGCGTGGGGCGGCGTGC GGCGCATTCGCTGAAGACCCCCGGGGCCCTGTTC 336
Qy	181 tctgcttctcctggcatccctctttgcggagcaaccatgtggaaaccccgctcttcgtcca 240
Db	337 TCTGCTTCTCTGGGCATCCCTTTGCGGAGCCACCATTGGGAACCCCGTCGCTTCTGCCCA 396
Qy	241 ccggaagcccaagcagccttgtcaggggtgtgtagacgctaacaacctccagagtgcttc 300
Db	397 CCGGAGCCCAAGCAGCCTTGCTCAGGGGTGTAGACGCTACAACCTTCAGAGAGTGTCTGC 456
Qy	301 taaccaatatgtgacacaccatacccaagttttgaggacaccgagatgtgnaaccccaac 360
Db	457 TTACCAATATGTGACACACCCCTATACCCAGGTTTTTGAGGGCACCGAGATGTGAACCCCAAC 516
Qy	361 cgtgagctgagcgagagctgctgtacctcaacgtygtgacacacatacccccgctaca 420

D	b	517	CGTAGCTGAGCAGGACTGCCTGTACCTCAACGTGTGACACACCATATACCCCGCCTTACA	576
O	Y	421	tcctccacccctgtccctcgtcttgatctatagtgggttgctctacagtgggcctctcc	480
		577	TCCCCAACCCCTGTCTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCTCC	636
O	Y	481	ttggaagtgtacgaatgaccgcttcttggtacagggccgagagactgtgtgtgtccatg	540
D	b	637	TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGACTGTGCTGTGTTCATG	696
O	Y	541	aactacccggtgtggagcctttggtctctctgcccctgcggggagccgagagccccgggc	600
D	b	697	AACTACCGGGGTGGAGCCTTTGGCTTCTGTGGCCCTGTCCGGGGAGCCGAGAGCCCCGGGC	756
O	Y	601	aatgtgtgtctcctgtaacagaggtctggccctgacgtgtgtcagagaaagtgtcagcc	660
D	b	757	AAATGTGGGTCTCTGTGATCAGAGGCTGGCCCTGCATGTGGTGCAGAGAACTGTGCAGCC	816
O	Y	661	ttcgggggtgaccgcgacatcagtgacgtgtttgggagagcgcgagaccgctcgtg	720
D	b	817	TTCCGGGGGTGACCCGACATCAGTGACGCTGTTGGGGAGAGCGCGGAGCCGCTCGGTG	876
O	Y	721	ggcatgcacctgtctcccgcccgagccgggctgttcacagggccgtgtcgcagagc	780
D	b	877	GGCATGCACCTGCTGTCTCCCGCCCAAGCCGGGCGCTGTCCACAGGGCCGTGCTCAGAGC	936
O	Y	781	gtgtcccccgaatgaaaccttggggccaagtgtggcattgtggagagagcccgctgcagggcca	840
D	b	937	GGTGCCCCCAATGGACCTTGGGGCCACAGGTGGGCATGGGAGAGGCCCGTCCGACGGCCACG	996
O	Y	841	caagctggcccaacctgttgggtctgtcctccaagcggcaatgttggaaatgacacagaactg	900
D	b	997	CAGCTGGCCCACTTGTGGGCTGTCTCTCAGCGCGCACTGTGGGAATGACACAGAGCTG	1056
O	Y	901	gtagcctgtccttcgaaacagaccagcgcaagtctcctgtgtgaaacagaaatgcagtgctg	960
D	b	1057	GTAAGCTTGCTTCGGACACGACCAAGCGCAGGTCTGTGTAAACAGAAATGGCACTGTCTG	1116
O	Y	961	ccctcaagaaagcgtcttcgggtctccttcgtctgtcgtgtgtagatgagaaacttcctcagt	1020
D	b	1117	CCTCAAGAAAGCGTCTTCCGGTTCCTCTCGTGCCGTGTGTAGATGGAGACTTCTCTCAGT	1176
O	Y	1021	gacacccccagagggccctcatcaacgcgggagagacttccaagggcctgcaggtgtctgtgg	1080
D	b	1177	GACACCCCCAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGTGGGT	1236
O	Y	1081	gtgtgtgaaagatgagggctcgtatttctgtgttlaagggccccaaggtctcagcaaaagac	1140
D	b	1237	GTTGGTGAAGGATGAGGGGCTCTATTCTTCTGTGTTAAGGGGGCCCCAGGCTTCAGCAAAAGAC	1296
O	Y	1141	aacgagttctctatcagccgggccgagttcctctgcccgggtgcgggtcgggttccccag	1200
D	b	1297	AACGAGTCTCTCATCAGCCGCGGCGAGTTCCTGGCCGGGGTGGGGTGGGGTTCCTCCAG	1356
O	Y	1201	gtaagtgaacctgacgcagcagcgtgtgttctctgcatlacacagactgtgtcatcccgag	1260
D	b	1357	GTAAGTGAACCTGGCAGCCGAGGCTGTGTCTCTGCATACACAGACTGCTGCATCCCGAG	1416
O	Y	1261	gaeccggaacgcctgagggggccctgagcgaatgtgtgtggggaaccacaatatgtcgtgtgc	1320
D	b	1417	GACCCGGCACGCGCTGAGGGAGGCCCCTGAGCGATGTGTGGGGCGAACACAAATGTCTGTGC	1476
O	Y	1321	ccccgtggcccaagctgtgtgtgtggcgaactgtgtcccaggtgtcccgggtctacgctcaagt	1380
D	b	1477	CCCCGTGGCCCAAGCTGGCTGGGGCGACTGGCTGCCCCAAGGTGCCCGGGTCTACGCTTACGTC	1536
O	Y	1381	tttgaacaacgctgtctccaagctctccttgccccctgttggattgggggtgtccccacaggttac	1440
D	b	1537	TTTGAACACCGTGTCTTCCACGCTCTCTGTGGCCCCCTGTGGATGGGGGTGCCCCACGCGCTAC	1596
O	Y	1441	gaagatcgaagttcatcttcttggatcccccttgagccccctctcgaaaactacacgacgaaggag	1500
D	b	1597	GAGATCGAGTTTCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCGAGAGAG	1656

QY 1501 aaatcttcgcccagcactgatcgatctatctgagccaactttgccgcacaggagatccc 1560
|||||
Db 1657 AAAATCTGCGCCAGCGACTGATGCGATACTGGGCCCACTTTGCCCGCACAGGGGATCCC 1716
QY 1561 aatgagcccgagagaccacaaggcccacaaatgccccctacacagcgagggtcacaag 1620
|||||
Db 1717 AATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCGTACACGGCGGGGCTCAGCAG 1776
QY 1621 taagttagtctgacactgagccgctgtagagtgcgagggggtctgcgagcccaagctgc 1680
|||||
Db 1777 TACGTTAGTCTGAGACCTGGCGCCGCTGAGAGTGGGGGGGCTGCGCGCCAGGCTGC 1836
QY 1681 gccctctggaaccgcttcctccccaattgtctcagcgccacc 1722
|||||
Db 1837 GCGTTCTGGAACCGCTCTCCCAATTTGCTCAGCGCCACCC 1878

RESULT 3
AR070205 AR070205 2256 bp DNA linear PAT 18-FEB-2000
LOCUS Sequence 5 from patent US 5891725.
DEFINITION AR070205
ACCESSION AR070205
VERSION AR070205.1 GI:7221093
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2256)
AUTHORS Soreq,H., Zakut,H. and Eckstein,F.
TITLE Synthetic antisense oligodeoxynucleotides and pharmaceutical compositions containing them
JOURNAL Patent: US 5891725-A 5 06-APR-1999;
FEATURES Location/Qualifiers
source 1..2256
BASE COUNT 390 a 757 c 680 g 429 t
ORIGIN

Query Match 99.8%; Score 1722; DB 6; Length 2256;
Best Local Similarity 100.0%; Pred.-No: 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggcccccagctgtctgtctgacacagcccttcctggtctcccaactccttcctc 60
|||||
Db 160 ATGAGGGCCCCCGAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCACTCTCTCTC 219
QY 61 ctccctctgctcctgagtgtagagagtgtaggggctcgaaggccggagagatgcagagctgtg 120
|||||
Db 220 CTCTCTGTGCTCTGGGTGAGGAGAGTGGGGGCTGAGGGCCGGAGAGATGCAGAGCTGCTG 279
QY 121 gtgacggtgcgtgagggccgagctgcgagcattcgcctgaagacccccgggggagcctgtc 180
|||||
Db 280 GTGACGGGTGCGTGGGGGGCGGCTGCGGCGCATTCGCTGAAGACCCCGGGGGCCCTGTC 339
QY 181 tctgcttctctgagcacccttctgcgagagccaccatgagcccccgctcgtctctgcga 240
|||||
Db 340 TCTGCTTCTCTGGGCAATCCCTTTGCGAGCCACCATGGGACCCCGTCTCTGCA 399
QY 241 ccgagagcccaagcagccttggtcagaggtgtgtagaagctacaaacttccagagtgtgc 300
|||||
Db 400 CCGGAGCCCCAAGCAGCTTGGTCAAGGGGTGTGAGACGCTTACAACCTTCCAGAGTGTCTGC 459
QY 301 taaccaatatgtgacacacctatacccaaggttttgaaggagcagagatgtgaaacccaac 360
|||||
Db 460 TACCAATATGTGACACCTATACCCAGGTTTGAAGGGCACCGAGATGTGAACCCCAAC 519
QY 361 cgtgagctgagcagagactgcctgtacactcaacgtgtgacacacatacccccgcgctaca 420
|||||
Db 520 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTTGTGACACCATATACCCCGGCTTACA 579
QY 421 tcccccaacctgtctcgtctgtgatactatgagggtggtctctacagtggggcctcctcc 480

|||||
Db 580 TCCCCCACCCCTGTCTCTGCTGTGATCTATGGGGGTGCTCTACAGTGGGGCTCTCC 639
QY 481 ttgagcgtgtacagtgcgccttctctgttacagcgacagagactgtgtgtgtcatg 540
|||||
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGACTGTGTGTGTTCATG 699
QY 541 aactaacgggtgtagagccttctgtctcctgtgcctgcgagggagccgagagccccgggc 600
|||||
Db 700 AACTACCGGGGTGGAGCCCTTTGGCTTCTCTGCGCCCTGCGGGGAGCCGAGAGGCCCGGCC 759
QY 601 aatgtggtctcctgtgatacagaggtgtgcctctgcagtgggtgcagagaaagctgcagcc 660
|||||
Db 760 AATGTGGGTCTCTGTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAACGTGGCAGCC 819
QY 661 ttcgggggtgacccgacatacagtgaagctgttttgaggagagcgcgagccgctcgtgtg 720
|||||
Db 820 TTCGGGGGTGACCCGACATCAGTGAAGCTGTGTTGGGGAGAGCGCGGGAGCCGCTCGGTG 879
QY 721 ggcatgacactgtgtcccccagcccgaggcctgttccaagggcgctgtcagagac 780
|||||
Db 880 GGCATGCACCTGTCTGTCCCGCCAGCCGGGGCTGTCCACAGGGCGCTGTCTGAGAGC 939
QY 781 ggtgcccccaatgaaaccttgagccacaggtgtggcatgagagagcccgctcagaggcagc 840
Db 940 GGTGCCCCCAATGAGCCCTGGGCCACAGGTGGGCATGGGAGAGGCCCGTCCGAGGGCCACG. 999
QY 841 cagctggcccaactgtgtgggtgtcctccagggcgacactgttggaaatgacacagagctg 900
|||||
Db 1000 CAGCTGGCCACACTGTGGGCTGTCTCCAGGGCGGCACTGGTGGGAATGACACAGAGCTG 1059
QY 901 gtagctgccttcgagacagacacagcgagcaggtctctgttgaaccaagaaatgcaagctgtg 960
|||||
Db 1060 GTAGCTGCTCTCGGACACGACACGAGCGGAGGCTCTGTGTAACCAACGAATGGCACGTGCTG 1119
QY 961 cctcaagaaagcgtctcctgggtctcctctgtgctgtgtgtagatgagagactcctcagt 1020
|||||
Db 1120 CCTCAAGAAAGCGTCTTCGGGTCTCTCTGTGCTGTGTGATGAGAGACTTCTCAGT 1179
QY 1021 gacacccagagggccctcaacacgcgagagacttccagggcctgcaggtgtcgtggt 1080
|||||
Db 1180 GACACCCAGAGGCCCTCTATCAACGCGGAGACTTCCAGGCTGAGTGTGTGGGT 1239
QY 1081 gtgtgaaagatgaggggtcgtatcttctgtgttacagggccccaggttcagcaaaagac 1140
|||||
Db 1240 GTGTGAAGATGAGGGGTGATTTTCTGTGTGTACGGGGCCCCAGGCTCAGCAAAAGAC 1299
QY 1141 aacgagctctcatcagccgggcccaggttcctgcgaggggtgcggtgtcgggttcccaag 1200
|||||
Db 1300 AAGAGTCTCTCATCAGCGGGGCCAGAGTCTCTGCGGGGTGCGGGTCTCCCAAG 1359
QY 1201 gtaagtacctggcagccgaggtgtgtgtcctgtcatcaacagagactgtgtcatcccgag 1260
|||||
Db 1360 GTAAGTGACCTGGCAGCGGAGGCTGTGGTCTCTCATTAACAGACTGGCTGCATCCCGAG 1419
QY 1261 gaaccggcagccttgaaggagggccctgaagcagatgtgtgtgagcaccacaatgtcgtgtgc 1320
|||||
Db 1420 GACCCGGCAGCCTGAGGAGGCCCCGTGAGCGATGTGTGGCGCACACAATGTGTGTGC 1479
QY 1321 cccgtggccagctgtgtgagcagactgtgcagaggtgtcccggtctcagcctaagctc 1380
|||||
Db 1480 CCCGTGGCCAGCTGTGGGTGGCGACTGTGCTGCCAGGGTGCCTGCTACGGCTACGTC 1539
QY 1381 ttgaaacacgtgtcttcaagctctcctgtgcctctgtgagtgtgggtgtcccaaggttac 1440
Db 1540 TTGGAACACCGTGTCTCAACGCTCTCTGCGCCCTGTGTGATGGGGGTGCCCAAGGCTAC 1599
QY 1441 gagatcgaattcatcttgggagatcccccttgagacccctctcgaactacacagcagaagag 1500
|||||
Db 1600 GAGATCGAGTTTATTTTGGGATCCCCCTGAGCCCTCTCGAAACTACACGGCAGAGAG 1659
QY 1501 aaatcttcgcccagcagactgatgcatactggccaactttgccgcacaggagatccc 1560
|||||

Db 1660 AAAATCTTGGCCCGACGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagcccccgaagaaccccaagggcccccataatggcccccgtacacggcggggtcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCCAATATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
QY 1621 taacttagctctggaacctgcgccgctggaaggtgcggcggggtctgcgcccagcctgc 1680
Db 1780 TACGTTAGTCTGGACCTGCGGGCGGTGAGGTGCGGGGGGCTGCGGCCAGGCGCTGC 1839
QY 1681 gcctctggaaccgcttcctccccaattgctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCTCCCAAAATGCTCAGCGCCACC 1881

RESULT 4

AR070207 3016 bp DNA linear PAT 18-FEB-2000
LOCUS AR070207
DEFINITION Sequence 7 from patent US 5891725.
ACCESSION AR070207
VERSION AR070207.1 GI:7221095
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3016)
AUTHORS Soreq, H., Zakut, H. and Eckstein, F.
TITLE Synthetic antisense oligodeoxynucleotides and pharmaceutical
JOURNAL Compositions containing them
Patent: US 5891725-A 7 06-APR-1999;
FEATURES Location/Qualifiers
Source 1..3016

BASE COUNT 497 a 1065 c 840 g 614 t
ORIGIN

Query Match 99.8%; Score 1722; DB 6; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggcccccgaagtctctgtcgacacgacctccctgctcccactccttcctc 60
Db 160 ATGAGGGCCCCGAGTGTCTGTCACACGCCCTTCCCTGGCTTCCCACTCCTTCTCCTC 219
QY 61 ctccctggtctcctggtgtgagaggtgggtctgagggccgggagatgacagctgtg 120
Db 220 CTCTCTGGCTCCTGCTGGTGAGAGAGTGGGGCTGAGGGCCGGAGAGATGCAGAGCTGTG 279
QY 121 gtgacgtgcgtgtggggcgcgctgcggggcaatcgcctgaagaaccccgggggcctgtc 180
Db 280 GTGACGGTGGCTGGGGCGGCTGCGGGGCATTCCGCTGAAGACCCCGGGGGCCCTGTC 339
QY 181 tctgcttctcctgggcatccctcttgcgagacccaatggaaccgcctgccttctgcca 240
Db 340 TCTGCTTCTCTGGGCATCCCTTTGCGAGCCACCCATGGACCCCGTCTTCTTGCCA 399
QY 241 ccggagcccaagcagcctgtgtcaggggtgtagacgtacaaacctccagagtgctgc 300
Db 400 CCGAGCCCAAGCAGCCTTGTGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtgacacccataatacaagggtttgagggcaccgagatgtgaaacccaac 360
Db 460 TACCAATATGTGACACCCCTATACCCAGGTTTGAAGGGCACCCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcgaaggaactgcctgtacctcaacgltgtggaacacatacccccggtctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCCTGTACTCAACGTGTGACACACCAATACCCCGGCTTACA 579
QY 421 tcccccaacctgtcctcgtctgatactatgggggtgtgcttctacagtggggcctcctcc 480
Db 580 TCCCCCACCCCTGTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCCCTCTCC 639

QY 481 ttgacgtgtacagatgagccgctctctgtgtacagggccgaagagactgtctgttcatg 540
Db 640 TTGACAGTGTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGGACTGTCTGTCTCATG 699
QY 541 aactaccgggtggagccttctgtctcctgcccctgcggggagccgaagagccccgggc 600
Db 700 AACTACCGGCTGGAGCCTTTGGCTTCTCTGCCCCCTGCCCCGGGAGCCGAGAGCCCCGGGC 759
QY 601 aatgtgtctccttgatcagaaggtgcccctgcagtgtgtgcagagaacgtgacagcc 660
Db 760 AATGTGGGTCTCTTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAACTGGCAGCC 819
QY 661 ttcggggtgacccgaacatcagtgacgtgtttggggagagcggcgagcccgctcgtgtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTGGGGAGAGCGCGGGAGCCCGCTCGGTG 879
QY 721 ggcattgacacctgtctgtcccgcccaagccggggcctgttccacaagggccggtgtcagagc 780
Db 880 GGCATGACACCTGTCTGCCCGCCCGCCAGCGCGGCTGTTCACAGGGCCGCTGTGCAGAGC 939
QY 781 ggtgcccccaatggaaccttgggccaaggtggcatgtggaagagccccgtgcagggccaag 840
Db 940 GGTGCCCCCAATGGAACCTTGCGCCACGGTGGCATGGAGAGGCCCGTGCAGAGGCCACG 999
QY 841 cagctggcccaaccttgtggtgtctcctcgaagcggaactgtgtggaatgacacagagctg 900
Db 1000 CAGCTGGCCCACTTGTGGGCTGTCTCTCCAGCGGCACACTGTGGGAATGACACAGAGCTG 1059
QY 901 gtagcctgtcctcggaacagacacagcagcgagctcgtgtgaacacagatgcaagctgtc 960
Db 1060 GTAGCCTGTCTCGGACACAGACACGAGCGCAGCTCTGTGAACACCAATGGCACGCTGCTG 1119
QY 961 cctcaagaagaagctcttcggtctcctcctcctcgtgtgtgtatgtgagaacttcctcagt 1020
Db 1120 CCTCAAGAAAGCTCTTCCGGTCTCTCTTCTGCTGTGTGATGAGACTTCTCTCAGT 1179
QY 1021 gacaccccaagggccctcatcaaacgcgggaactccaagcgctgcaggtgtgtgtgt 1080
Db 1180 GACACCCCAAGGCCCTCATCAACCGGGCGAGTCTCTGGCCGGGTGCGGGGTCTCCCCAG 1239
QY 1081 gtgtgaagatgagggctcgtatcttctgttcaagggccccagagcttcagcaagac 1140
Db 1240 GTGTGAAGATGAGGGCTCGTATTCTGTGTACGGGGCCCCAGGCTTCAACAAGAC 1299
QY 1141 aacgagctctcatcagccggcggaagcttcctgcggcggtgtcggtgtcccccag 1200
Db 1300 AACGAGTCTCTCATCAACCGGGCGAGTCTCTGGCCGGGTGCGGGGTCTCCCCAG 1359
QY 1201 gtaagtacacctggcagccgagagctgtgtcctgtcatcaacagactgtctcatcccgag 1260
Db 1360 GTAAGTACCTGGCAGCCGAGGCTGTGTCTGTCAATTACACAGACTGTGCTCATCCCGAG 1419
QY 1261 gaccggcagcgtctgagggagggccctgagcgaatgtgtgtggcgacaacaatgtcgtgtgc 1320
Db 1420 GACCCGGCAGGCTGAGGGAGGCGCTGAGCGATGTGTGGCGCACAAATGTCTGTGTGC 1479
QY 1321 ccggtggcccaagctgtgtggcgactgtgcccagaggtgcccgggtctacgacctacgtc 1380
Db 1480 CCGGTGGCCCAAGCTGTGGCGACTGTGCGCAAGGTGCCCCGGGTCTACGCTTACGCTC 1539
QY 1381 ttgaaacacggtcttccaagctctcctgtgcccctgtgatagggggtgtccccaaggtctac 1440
Db 1540 TTTGAACACCGGTCTTCCACGCTCTCTGTGGCCCTGTGATGGGGGTGCCACAGGCTAC 1599
QY 1441 gagatcgagttcatcttgggatacccccctggaaacccctctcgaaactacaacggcagagag 1500
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaaactctgcaccaagcgaactgtatgatacttggccaacttgcgcgcacaaaggagatccc 1560
Db 1660 AAAATCTTCCGCCACGCACTGATGCGATGTGGGCCAACTTTGCCCCGACAGGGGATCCC 1719
QY 1561 aatgagcccgaggaaccccaagggccccaacatggccccgtacacagcggggtctcagcag 1620

Db	1720	AATGAGCCCCGAGACCCCAAGGCCCCACAATGGCCCCGTACACGGCGGGGCTCAGCAG	1779
QY	1621	tacgttagtctcgaacctgcgcccgtctgagatlgcgcgggggtcgcgccagccctgc	1680
Db	1780	TACGTTAGTCTGGAACCTCGCGCCCGCTGGAGGTGC GGCGGGGGCTGCGGCCAAGGCTGC	1839
QY	1681	gccttcctgaaaccgcttcctcccccaaatctgctcagcgccacc	1722
Db	1840	GCCCTTCGGAACCGCTTCCTCCCAATTGCTCAGCGGCACC	1881

RESULT	5
1000000	

LOCUS	AR070206	3096 bp	DNA	linear	PAT 18-FEB-2000
DEFINITION	Sequence 6 from patent US 5891725.				
ACCESSION	AR070206				
VERSION	AR070206.1	GI:7221094			

SOURCE ORGANISM	Unknown.
	Unknown.

Unclassified.
REFERENCE 1 (bases 1 to 3096)

AUTHORS	Soreq, H., Zakut, H. and Eckstein, F.
TITLE	Synthetic antisense oligodeoxynucleotides and pharmaceutical

compositions containing citrulline
Patent: US 5891725-A 6 06-APR-1999;
JOURNAL

FEEDONES	LOCATION/QUALITIES
SOURCE	1. . 3096

BASE COUNT	509 a	1089 c	872 g	626 t
------------	-------	--------	-------	-------

ORIGIN

Query Match	99.8%	Score 1722;	DB 6;	Length 3096;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1722: Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

QY	1	atgaggcccccgagtgctgtctgcacacgacctccctggtctcccaactctctcc	60
Db	160	ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCACTCTTCTCCTC	219
QY	61	ctccctgtgctcctctggtgtgagagtggtgggtctgaggccggagagatgcagagctgtg	120
Db	220	CTCCTCTGGCTCTCTGGGTGAGAGACTGGGGGCTGAGGGCCGGAGAGATGCAGACTGCTG	279
QY	121	gtgacggtgctgtgtgtggccggtctgcgtgggcatctgcctgaaagacccccggggccctgtc	180
Db	280	GTCAGCGGTGCGTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCCGGGGGCCCTGTGTC	339
QY	181	tctgtcttctcctgtggcatccccccttgcgtgagccaccccatgtgagccccgtctgccttctgcca	240
Db	340	TCTGCTTTCCTGGGCATCCCTTTGGCGAGCCACCACCATGGACCCCGTCGCTTCTGCCA	399
QY	241	ccggagccccaagcagccttgtcaagggtgtgtagcgtacaaacctccagagtgtctgc	300
Db	400	CCGGAGCCCAAGCAGCCTTGGTCAAGGGGTGTTAGACGCTTACAACCTTCCAGAGTGTCTGC	459
QY	301	taccaatatgttgaaacaccctatacccaagttttgaggcaccgagatgtgaaaccccaac	360
Db	460	TACCAATATGTGGACACCCCTATACCAAGTTTGAAGGGCACCGAGATGTGAACCCCAAC	519
QY	361	cgtgagctgagcgagagactgctgttaacctcaacgltgtgacacataccccggtctaca	420
Db	520	CGTGAGCTGAGCGAGGACTGCGCTGACTCAACGTGTGACACACCATACCCCCGGCCTACA	579
QY	421	tcccccaacccctgtcctcgtctgtgatactatgtgggtgtgcttctacagtgtgggtctctcc	480
Db	580	TCCCCCAACCCCTGTCTCTCGTGTGATCTATGGGGGTGGCTTCTTACAGTGGGGCCTCTCC	639
QY	481	ttggagcgtgtacgattggtgcgctctcttctgtgtacaggtccgagagagactgtgtctgttccatg	540
Db	640	TTGGAGCTGTACGATGTGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGCTGTCCATG	699

QY	541	aactaacccggtgagaccttctggcttccctgccccgagccgagagcccccggc	600
Db	700	aaactaacccgggtggagccctttggccttccctggccccctgccccggagccgagagcccccggc	759
QY	601	aatgtgggtctcctggtacgaagcgtgccccctgcagtgggtgcagagagaacgtgcagcc	660
Db	760	aatgtgggtctcctggtacgaagcgtgccccctgcagtgggtgcagagagaacgtgcagcc	819
QY	661	ttcgggggtgacccgaatcagtgaacgtgttctggggagagcgcggaagccgctcgtg	720
Db	820	ttcgggggtgacccgaatcagtgaacgtgttctggggagagccgggagccgctcggctg	879
QY	721	ggcatgcacctgtgtccccgccccagccggggccctgtttccacagggccgtgtcagaagc	780
Db	880	ggcatgcacctgtgtccccgccccagccggggccctgtttccacagggccgtgtcagaagc	939
QY	781	ggtgcccccaatggaacctgggccaaggttggtgatggagagagcccgctgcagagccacg	840
Db	940	ggtgcccccaatggaacctgggccaaggttggtgatggagagagcccgctgcagagccacg	999
QY	841	caagctggcccaacctgtgtggcgtgtcctccagggcgcactgttggaatgacacagaagctg	900
Db	1000	caagctggcccaacctgtgtggcgtgtcctccagggcgcactgttggaatgacacagaagctg	1059
QY	901	gtagcctgccttcggaacgaaccaagcgcaagcttcctgtgtgaaccaagaaatgcagtgctg	960
Db	1060	gtagcctgccttcggaacgaaccaagcgcaagcttcctgtgtgaaccaagaaatgcagtgctg	1119
QY	961	ccctcaagaaagcgtcttccggtctccttcgtgcctgtgtagatgagaccttcacgt	1020
Db	1120	ccctcaagaaagcgtcttccggtctccttcgtgcctgtgtagatgagaccttcacgt	1179
QY	1021	gacacccccagagccctcatcaacgcgggagacctccaagcctgcaggtgtcgtgggt	1080
Db	1180	gacacccccagagccctcatcaacgcgggagacctccaagcctgcaggtgtcgtgggt	1239
QY	1081	gtgtgtaaagatatgaggctcgtaatttctggtttaaggggccccagccttcagcaaaagac	1140
Db	1240	gtgtgtaaagatatgaggctcgtaatttctggtttaaggggccccagccttcagcaaaagac	1299
QY	1141	aacgagctctcatcagccgggcccagttccttgcccgggtgcgggtcggggttccccag	1200
Db	1300	aacgagctctcatcagccgggcccagttccttgcccgggtgcgggtcggggttccccag	1359
QY	1201	gtaagtgaacctgggcagccgagcgtgtgtctctgcatcaacagacctgctgcataccgag	1260
Db	1360	gtaagtgaacctgggcagccgagcgtgtgtctctgcatcaacagacctgctgcataccgag	1419
QY	1261	gaacccggcacgcctgagggagggccctgagcgaatgtgtgtggcgacacacaatgtcgtgtgc	1320
Db	1420	gaacccggcacgcctgagggagggccctgagcgaatgtgtgtggcgacacacaatgtcgtgtgc	1479
QY	1321	cccgtggccccagctggtcgtggcgactgtgctgcaccaggggtgcccgggtctacgcctacgtc	1380
Db	1480	cccgtggccccagctggtcgtggcgactgtgctgcaccaggggtgcccgggtctacgcctacgtc	1539
QY	1381	tttgaacaaccggtgtctccaagctctccttgccccctgtgtgagtgggggtgcccccaagcgtac	1440
Db	1540	tttgaacaaccggtgtctccaagctctccttgccccctgtgtgagtgggggtgcccccaagcgtac	1599
QY	1441	gagatcgaattcatcttgggataccccctggaacccctctcgaactacacagcgcagaggag	1500
Db	1600	gagatcgaattcatcttgggataccccctggaacccctctcgaactacacagcgcagaggag	1659
QY	1501	aaaaatctgcgccagcgactgtatgcatalactgggcgaacttgcccgcacaggggatccc	1560
Db	1660	aaaaatctgcgccagcgactgtatgcatalactgggcgaacttgcccgcacaggggatccc	1719
QY	1561	aatgaagccccgagaccccccaaggcccccaaatgtgcccccgtaacagcgcggggtcgaagcag	1620
Db	1720	aatgaagccccgagaccccccaaggcccccaaatgtgcccccgtaacagcgcggggtcgaagcag	1779

QY 1621 tacgttagctgacactgcgcgcgtgaggtgcggcggtgctgcgcgccagcctgc 1680
|||||
Db 1780 TACGTTAGTCTGGACCTGGCGGCTGAGAGTGGCGGGGCTGGCGCCAGGCTGC 1839
QY 1681 gcctctggaaccgcttcctcccaattgctcagcgcacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCCCTCCCAATGCTCAGCGGCACC 1881

RESULT 6
AX275254 5767 bp DNA linear PAT 29-OCT-2001
LOCUS Sequence 3 from Patent WO0171014.
DEFINITION AX275254
ACCESSION AX275254
VERSION AX275254.1 GI:16547674
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (sites)
AUTHORS Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants
JOURNAL Patent: WO 0171014-A 3 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES
source location/Qualifiers
1..5767
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM034"

BASE COUNT 1390 a 1521 c 1495 g 1361 t
ORIGIN

Query Match 99.2% Score 1712; DB 6; Length 5767;
Best Local Similarity 100.0% Pred. No. 0;
Matches 1712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ccccgagctgtctgtgcacagcctccctgctcccaactctctctctc 66
|||||
Db 844 CCCCCAGTGTCTGCTGCACAGCCTTCCCTGGCTCCCACTCTCTCTCTC 903

QY 67 tggctctggtgtgagaggtggtggtgagggcgagagatgacagctgtgtgacg 126
|||||
Db 904 TGGCTCTGGGTGAGAGAGTGGGGCTGAGGGCGGAGATGACAGCTGTGTGACG 963

QY 127 gtcgctggggcggtgcgtgcagcattgcctgaagaccccggtgctgtctgt 186
|||||
Db 964 GTGCGTGGGGCGGCTGGCGGCACTTGCCTGAAGACCCCGGGGCTGTCTGTCT 1023

QY 187 ttcctggcatccctctgagcagccacccatgagacccctgctctctgacacggag 246
|||||
Db 1024 TTCTTGGGCATCCCTTTGGGAGCCACCCATGGGACCCCGCTTCTGCCACCGAG 1083

QY 247 cccaagcagcctgtgtcaggggtgtagagctacaacctccagagtgtctgtaccaa 306
|||||
Db 1084 CCCAAGCAGCCTTGTGTAGGGGTGAGACGTACAACTTCCAGAGTGTGTGTACCAA 1143

QY 307 tatgtgacacccatacccaagttttgagggacccagatgtgaaaccccaacgtgag 366
|||||
Db 1144 TATGTGACACCCATATCCCAAGTTTGAAGGGACCCGAGATGTGAAACCCCAACCTGAG 1203

QY 367 ctgagcagagactgctgtactcaacggtgtgacacacatacccccggtacatacccc 426
|||||
Db 1204 CTGAGCGAGGACTGCTGTACTCAACGTGTGACACCACTAACCCCGGCTTACATCCCCC 1263

QY 427 accctgtcctcgtctgatactatgggggtgtgctctacagtgggctcctccttgag 486
|||||
Db 1264 ACCCTGTCTCTGCTGTGATATAGGGGTGTGCTTACAGTGGGGCTCTCTGTGAGC 1323

QY 487 gtgtacgaltgcccgtctctgtgtacagcgccgagagactgtgtgtgtcatgaactac 546

Db 1324 GTGTACGATGGCCGCTTCTTGTGTACAGCCGAGAGGACTGTGTGTTCATGAATAC 1383
QY 547 cgggtggagccttltgctcctgagccctgcgggagccgagagcccggtgcaatgtg 606
|||||
Db 1384 CGGTTGGAGCCTTTGGCTTCTGTGCTGCCCCGGGAGCCGAGAGGCCCGGCAATGTG 1443

QY 607 gttctctgatacagaggtgtgacctgagtggtgtcagagaaagtgtgacccctcggg 666
|||||
Db 1444 GGTCTCTGTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAAAGTGGCAGCCTCGGG 1503

QY 667 gttgacccgacatcagtgacgctgtttggtggagagcgcggtgagccctcgtgtggcatg 726
|||||
Db 1504 GGTGACCCGACATCAGTGACGCTGTTGGGGAGAGCGCGGGAGCCGCTCGGTGGCATG 1563

QY 727 caactgtgtcccccagcccggtgctgttccacagggccgtgtcagagcggtgtcc 786
|||||
Db 1564 CACTGTGTGTCCTCCCGCCAGCCGGGGCTGTTCACAGGCGCGTGTGCAGAGCGGTGCC 1623

QY 787 cccaatgacccctgggccaaggtgtggtgagagagcccgctgcagggccaagcagctg 846
|||||
Db 1624 CCCAATGGAACCTGGGGCCACGGTGGGCAATGGGAAGGCCCGTGCAGGGCCACGCACTG 1683

QY 847 gccacactgtgggtgtctctccagggcgactgtgtggaatgacacagactgtgtagcc 906
|||||
Db 1684 GCCACACTGTGGGCTGTCTCTCCAGCGCGGCACTGTGGGAATGACACAGACTGTGTAGCC 1743

QY 907 tgccttcggaacacgacacagcgaggtctgtgtgaaacacgaatgtgcagtgtcctcaa 966
|||||
Db 1744 TGCTTGGACACGACCAAGCGGAGGTCTGTGTGAACCAAGAAATGGACGTGCTCTCAA 1803

QY 967 gaaagcgtctccggttctcctcgtgtcgtgtgtgatatgagactcctcagtacacc 1026
|||||
Db 1804 GAAAGCGTCTCCGGTCTCTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1863

QY 1027 ccagagccctcatcaacgcgagagactccaagcctgcaggtgtgtgtgtgtgtgtgt 1086
|||||
Db 1864 CCAGAGCCCTCATCAACGCGGAGACTTCCACGGCCTGCAGAGTGTGTGTGTGTGTGTGT 1923

QY 1087 aagatgaggtctglatlctgtgttaccggggcccaagcttcaagcaaacagagag 1146
|||||
Db 1924 AAGGATGAGGGCTCGTATTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1983

QY 1147 tctctcatcagccgggcccaggttctctgcccgggtgtcgggttccccaagttaagt 1206
|||||
Db 1984 TCTCTCATCAGCCGGGGCCGAGTTCCTGGCGGGGTGCGGGTCTCCCAAGTAAGT 2043

QY 1207 gacctgtcagccgaggtgtgttctctgattacaacagactgtgtcatcccgaggaaccg 1266
|||||
Db 2044 GACCTGGCAGCCGAGGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2103

QY 1267 gcaagcctgaagggaagccctgagcagatgtgtgtgacacacaaatgtgtgtgtgtgt 1326
|||||
Db 2104 GCACGCTGAGGGAGGCGCTGTGAGCGATGTGTGGGCGACCAAAATGTGTGTGTGTGTGT 2163

QY 1327 gcccaagctgtgtggtgagactgtgtcccaagggtgtccgggtctacgctactgtgtgt 1386
|||||
Db 2164 GCCCAGCTGGCTGGCGACTGGCTGCCCAAGGTGCCGGGTCTACGCTTGTGAA 2223

QY 1387 caacgtgttccaaagctctcctgtgcccctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1446
|||||
Db 2224 CACGTTGCTTCCACGCTCTCTGCCCCCTGTGATGGGGGTGCCCAAGGCTTACGAGATC 2283

QY 1447 gaattcatcttgggatacccccctgtgaacccctctgaaactacaacgagagagaaatc 1506
|||||
Db 2284 GAGTTCATCTTGGGATCCCGCTGGAACCCCTCTCGAAACTACAGGAGAGAGAAATC 2343

QY 1507 ttcgcccagcagactgtatgcatactgtgtgcaacttgcgcgcagaggggtatcccaatgag 1566
|||||
Db 2344 TTCGCCACGCACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCCAATGAG 2403

QY 1567 cccgagagccccaaaggccccaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1626
|||||

Db 2404 CCCGAGACCCCAAGGCCCCACAAATGCCCCCTACACGGCGGGGCTCAGCAGTACGTT 2463
QY 1627 agctgacccgcgcccgtgaggtgagggcggggtgctgcgcgcgcagcctgccttc 1686
Db 2464 AGTCTGAGACTGCGGGCGCTGAGAGTGGCGGGGCTGCGCGCCAGGCCCTGCGCCTTC 2523
QY 1687 tggaaacgcttcctcccaaatgtcagcgc 1718
Db 2524 TGGAAACCGCTTCTCCCAATTGTCTACGCC 2555

644-2555 / 803

RESULT 7
AX275255 14446 bp DNA linear PAT 29-OCT-2001
LOCUS AX275255 Sequence 4 from Patent WO0171014.
DEFINITION AX275255
ACCESSION AX275255
VERSION AX275255 GI:16547675

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (sites)
AUTHORS Mor,T., Soreq,H., Arentzen,C. and Mason,H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants

JOURNAL
Patent: WO 0171014-A 4 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; MOR,
Tsafir (US) ; Soreq, Hermona (IL) ; Arentzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES
source 1.14446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM036"

BASE COUNT 3231 a 3831 c 4046 g 3042 t 296 others
ORIGIN

Query Match 99.2%; Score 1712; DB 6; Length 14446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cccccgagtgtctgtgcacacgccttcctgcttcccaactcctctcctc 66
Db 148 CCCCCGAGTCTGCTGCACACGCCCTCCCTGGCTCCCACTCCTCTCCTC 207
QY 67 tggctcctggttgagagtggtgagggcgggagagagatgcagagctgtgtgacg 126
Db 208 TGGCTCTGGGTGAGAGAGTGGGGGCTGAGGGCCGGAGGATGCAGACTGTGTGACG 267
QY 127 gtgcgtggggcgcgctgcgggcatcgcctgaagaccccggggcccctgtctgt 186
Db 268 GTGCGTGGGGCGCGCTGCGGGCATTCGCTGAAGACCCCGGGGCGCTGTCTGCT 327
QY 187 ttcctggcatcccttgcgagaccccaatggaaccccgctgccttctgcacagag 246
Db 328 TTCTGTGGGATCCCTTTGGGAGCCACCATGGACCCCGTGTCTGTCCACCGGAG 387
QY 247 cccaagcagccttgatcaggggtgtgacgctacaaccttccagagtgtgtacc 306
Db 388 CCCAAGCAGCCTTGGTCAAGGCTGTGACGCTTACACCTTCCAGAGTGTGTACCAA 447
QY 307 tatgtgacacccatataccaggtttgagggcacccagagatgtgaaaccccaacg 366
Db 448 TATGTGACACCTATACCAAGTTTGAAGGGCACCGAGATGTGAACCCCAACCGTAG 507
QY 367 ctgagcagagagctgctgtacctaagctgtgacacacataccccggcctacatcccc 426
Db 508 CTGAGCGAGAGACTGCTGTACCTCAACGTGTGACACCATACCCCGGCTTACATCCCC 567
QY 427 acccctgtcctcgtctgtatctatgggggtgctctacagtgggcctcctccttgac 486
Db 568 ACCCGTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCTCTCTTGAC 627

QY 487 gtgtacagatgcccgttcttgtgtacagggccgagagagactgtgtgtgtgtccatgaactac 546
Db 628 GTGTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGACTGTGTGTGTGTATGAATAC 687
QY 547 cgggtggagaccttggcttctcctgtgcccctgcggggagccgagagagcccgggcaatgtg 606
Db 688 CGGGTGGAGCCTTGTGGCTTCTGTGGCCCTGCGCGGGAGCCGAGAGGCCCGGCAATGTG 747
QY 607 ggtctcctgtgacagagaggtgcccctgcagtggtgtgagagagagctgtgagccttcggg 666
Db 748 GGTCTCTGTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAGAACGTGGCAGCCTTCGG 807
QY 667 ggtgacccgacatcagtgacgtgttggggagagagcgggagagccgctcgtgtggacatg 726
Db 808 GGTGACCCGACATCAGTGACGCTGTTGGGGAGAGCCGGGAGCCGCTCGTGGGATG 867
QY 727 caccctgttcccccgcccaagcggggctgttccacagggcgctgtcagagcggtgcc 786
Db 868 CACCTGTGTCCCGCCCAAGCCGGGGCTGTTCACAGGGCCGTGCTGCAGAGCGGTGCC 927
QY 787 cccaatgacccctgggccaagtggtgcatggagagagccgctgcagggccaagcagctg 846
Db 928 CCAATGACCCCTGGGCCCAAGGTGGGCAAGAGGCCCGCTGCAGAGGCCACGCACTG 987
QY 847 gccacactgtgggtgtctcctcagggcgagctgtggagatgacacagagctgtgtacc 906
Db 988 GCCACACTGTGGGCTGTCTCCAGGGCGCACTGTGGATGACACAGAGCTGTGTAGCC 1047
QY 907 tgccttcgacacagcagcgagctcctgtgtgaaccaagatgagctgtgtcctcaa 966
Db 1048 TGCTTCGAGACAGACAGCAGCGCAGGTCTGTGTGAACACAGATGGCAGCTGTGCTCAA 1107
QY 967 gaaagcgtcttcgggttctcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1026
Db 1108 GAAAGCGCTTTCGGTCTCTCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1167
QY 1027 ccagagccctcatcaaacgcgaggagacttccacggcctgcaggtgtgtgtgtgtgtgt 1086
Db 1168 CCAGAGGCCCTCATCAACGCGGGAGACTTCCAGGCCCTGCAGGTGTGTGTGTGTGTGT 1227
QY 1087 aagatgagggctcgtatcttctgtgttaccggggcccaagcttcagcaabaagacaagag 1146
Db 1228 AAGATGAGGGCTGTATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1287
QY 1147 tctctcatcagccgggcccaggttccctggtgcgggggtgcgggtgcgggttcccaagtgat 1206
Db 1288 TCTCTCATCAGCCGGGCGAGTTCCTGCGCGGGGTGCGGGTCCCAAGGTAAGT 1347
QY 1207 gacctgagcagcaggtgtgtctctgcattacacagactgtgtcatcccgagagcccg 1266
Db 1348 GACCTGGCAGCCGAGGCTGTGTCTCTGCATTAACAGACTGGCTGCATCCGAGGACCCG 1407
QY 1267 gcaagcctgagggagggccctgagcagatgtgtgtggcgacacacaatgtcgtgtgcccg 1326
Db 1408 GCACGCCCTGAGGAGGCCCTGAGCGATGTGTGGCGGACCAACAATGTCTGTGCCCGTG 1467
QY 1327 gccagctgtgtggcgactgtgtcccaagggtgccgggtctacgcctacgctccttgaa 1386
Db 1468 GCCAGCTGGCTGGGCGACTGTGCTGCCCAAGGTGCCCGGGTCTACGCTCTTTGAA 1527
QY 1387 caccgtgttccacagctcctcctggccccctgtgagtgggtgtgcccaagcgtacagatc 1446
Db 1528 CACCGTGTTCACAGCTCTCTGCGCCCTGTGTGATGGGGTGGCCCAAGGCTACGAGATC 1587
QY 1447 gagtcatctttgggatcccccttgagccctctcgaabaactacacggagagagaatc 1506
Db 1588 GAGTTCATCTTTGGGATCCCCCTGAGACCCCTCGAAACTACACGGCAGAGAGAAATC 1647
QY 1507 ttgcaccaagcagatgtgatactgtggccaacttgcggcacaagggatcccaatgag 1566
Db 1648 TTGCCCCAGCGACTGATGCGATCTGGGCCCACTTTGCCCGCACAGGGGATCCCAATGAG 1707

QY 1567 ccccgagaccacagagcccaatgccccgtacacggcggggtcagcagtc 1626
Db 1708 ccccgagaccacagagcccaatgccccgtacacggcggggtcagcagtc 1767
QY 1627 agctggacctgcygcgcgtgaggtgcygcggcgctgcgcacagcctgcctc 1686
Db 1768 agctggacctgcygcgcgtgaggtgcygcggcgctgcgcacagcctgcctc 1827
QY 1687 tggaaaccgcttcctccccaattgctcagcgc 1718
Db 1828 TGGAAACCGCTTCTCTCCCAATGCTCAGCGC 1859
RESULT 8
LOCUS HUMACHE01 1355 bp DNA linear PRI 14-JUN-1995
DEFINITION Human acetylcholinesterase (ACHE) gene, exon 2.
ACCESSION L22559
VERSION L22559.1 GI:853589
KEYWORDS acetylcholinesterase.
SEGMENT 1 of 4
SOURCE Homo sapiens blood DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Soreq, H.E., Ben-Aziz, R., Prody, C.A., Seidman, S., Gnatt, A.,
Neville, L., Lleman-Hurwitz, J., Lev-Lehman, E., Ginzberg, D.,
Lapidot-Lifson, Y. and Zakut, H.
Molecular cloning and construction of the coding region for human
acetylcholinesterase reveals a G + C-rich attenuating structure
Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9688-9692 (1990)
9108577
JOURNAL MEDLINE
REFERENCE 2 (sites)
AUTHORS Li, Y., Camp, S., Rachinsky, T.L., Getman, D. and Taylor, P.
Gene structure of mammalian acetylcholinesterase. Alternative exons
dictate tissue-specific expression
J. Biol. Chem. 266 (34), 23083-23090 (1991)
92078174
JOURNAL MEDLINE
REFERENCE 3 (bases 1 to 1355)
AUTHORS Bartels, C.F., Zelinski, T. and Lockridge, O.
Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene
accounts for YT blood group polymorphism
Am. J. Hum. Genet. 52 (5), 928-936 (1993)
93256075
JOURNAL MEDLINE
COMMENT On Jun 8, 1995 this sequence version replaced gi:348159.
**Reference [1] reports bases 1-1224;
Reference [J. Biol. Chem. 266, 23083-23090 (1991)] reports bases
1225-1233;
Reference [3] reports bases 1234-1358**.
FEATURES
source location/Qualifiers
1. 1355
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7q22"
/tissue_type="blood"
<1. 136
/gene="ACHE"
/note="G00-118-746"
/number=1
137. 1224
/gene="ACHE"
/note="G00-118-746"
/number=2
1213
variation /gene="ACHE"
/note="C [His(YT1)] to A [Asn(YT2)], YT blood group
polymorphism"
/replace="a"
BASE COUNT 207 a 462 c 416 g 270 t
ORIGIN

Query Match 62.0%; Score 1070; DB 9; Length 1355;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagcccccgagtgctgtgtgacacgcttcctgtgtccctcctcctc 60
Db 157 ATGAGCCCCCGAGTGTGTGTGTGACACGCTTCCCTGGCTTCCCACTCTCTCTC 216
QY 61 ctccctgtgtcctgtgtgtgagagtggtggtcagagggcgagagatgcagagctgtg 120
Db 217 CTCTCTGTGCTCTCTGTGGGTGAGAGAGTGGGGCTGAGGGCCGGAGAGATGACAGCTGCTG 276
QY 121 gtgacggtgctgtggggcgcgctgcggggcaltcgcctgaagacccccggggccctgtc 180
Db 277 GTGACGGGTGGGTGGGGCGCGCTGGGGGCATTCGCCGTGAAGACCCCCGGGGCCCTGTC 336
QY 181 tctgtcttcctgtggcatcccttgcggagccaccatggaaccccgctccttctgcca 240
Db 337 TCTGCTTCTCTGGGCATCCCTTTGGGAGCCACCCATGGGACCCCGCTTCTGTGCA 396
QY 241 ccgagccccagcagcctgtgtcaggggtgtgtagcgtctacacaccttcagagtgctgc 300
Db 397 CCGAGCCCAAGCAGCCTTGTGTACGGGGTGTAGACGCTTACACCTTCCAGAGTGTCTGC 456
QY 301 taccaatatgtgacacacctataccaggltttagggcaccagagatgtgaaccccaac 360
Db 457 TACCAATATGTGACACCCCTATACCCAGTTTTGAAGGACACCGAGATGTGGAACCCCAAC 516
QY 361 cgtgagctgagcagagagctgcctgtacctcaacggtgtgacacacacaccccgccctaca 420
Db 517 CGTGAGCTGAGCGAGAGACTGCTGTACCTCAACGTGTGACACCAATACCCCGGCTACA 576
QY 421 tccccccacctgtcctcgtctgtgatatgtgggggtgtgcttctacagtgggctcctcc 480
Db 577 TCCCCCACCCCTGTCTCTGTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCCCTCTCC 636
QY 481 ttgagcgtgtagatgagcgccttctgtgtacagggcgagagagagctgtgtgtccatg 540
Db 637 TTGGACGTGTACGATGGCCGCTTGTGTGTACAGGCCGAGAGGAGACTGTGTGTCCATG 696
QY 541 aactaccgggtggagccttgtgtcttcctgtgcctgcggggagccgagagggcccggc 600
Db 697 AACTACCGGGGTGGAGCCTTGTGTCTTGTGCCCCCTGCGGGGAGCGAGAGGCCCGCGGC 756
QY 601 aatgtggtctcctgtgatcagaggtgcctcgtcagtggtgcagagagagctgtgcagcc 660
Db 757 AATGTGGGTCTCTGTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAGAGCTGCCAGCC 816
QY 661 ttccggggtgaccccgacatcagtgtgctgttggggagagcgggagccgcctcgtgtg 720
Db 817 TTCGGGGGTGACCCGACATCAGTACGCTGTTGGGGAGAGCGCGGAGACCCGCTCGGTG 876
QY 721 ggcattgacatgtgttcccccggccggggcctgttccacagggccgtgtcagagc 780
Db 877 GGCATGACATGCTGTCTCCCGCCCGCCAGCGGGCCGTGTCCACAGGGCCGTGTGACAGC 936
QY 781 ggtgcccccaatggaacctgtggccaaggtgtgcatgtggaagagcccgctgcagggccag 840
Db 937 GGTGCCCCCAATGAGCCCTGGGGCCACGCTGGGCATGGGAGAGGCCCGCTGCAGGGCCACG 996
QY 841 cagctgccccacatgtgtgtgtgtcctccagggcgagcactgtgtggaatgacacagagctg 900
Db 997 CAGCTGGCCACCTTGTGGGTGTCTCTCCAGCGGCGCACTGTGGGATGACACAGAGCTG 1056
QY 901 gtagcctgcttcgagacagcagcgaggtcctgtgtgaacacagaaatgacagctgtgtg 960
Db 1057 GTAGCCTGCTTCGGACACGACGAGCGCAGGTCTGTGTAACCAAGATGGCAGCTGCTG 1116
QY 961 cctcaagaaagcgtcttcgcgttctcctcgtgtcctgtgtgtgtagatgtagacttcctcagt 1020
Db 1117 CCTCAAGAAAGCGTCTTCGGGTCTCTCTCTGCTGCTGTGTGTGATGAGACTTCTCTCAGT 1176
QY 1021 gacaccccaagagccctcatcaacgcgggagacttccacggtcctgcaggt 1070

```
|||||
Db 1177 GACACCCAGAGGCCCTCATCAACGGCGGAGACTTCACAGCGCTGCAAGT 1226

RESULT 9
HUMACHEB 4185 bp DNA linear PRI 23-JUN-1995
LOCUS Homo sapiens acetylcholinesterase (ACHE) gene, exons 2-6.
DEFINITION L42812
ACCESSION L42812
VERSION L42812.1 GI:854682
KEYWORDS acetylcholinesterase.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Bartels,C.F., Zelinski,T. and Lockridge,O.
TITLE Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene
JOURNAL accounts for YT blood group polymorphism
MEDLINE Am. J. Hum. Genet. 52 (5): 928-936 (1993)
93256075
REFERENCE 2 (bases 1 to 4185)
AUTHORS Bartels,C.F., Moriarty,P.L., Becker,R.E., Robbs,R.S.,
TITLE Sorenson,R.C., Mountjoy,C.P. and Lockridge,O.
JOURNAL Polymorphic sites in the acetylcholinesterase gene of patients with
FEATURES Alzheimer's disease
SOURCE Unpublished (1995)
location/Qualifiers
1.4185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7q22"
/cell_type="white blood cell"
75.1162
/gene="ACHE"
/note="G00-118-746"
/number=2
/evidence=experimental
75.4185
/gene="ACHE"
1151
/gene="ACHE"
/note="YT blood group; His322/Asn"
/replacement="a"
1163.1508
/gene="ACHE"
/note="G00-118-746"
/number=2
/evidence=experimental
1509.1993
/gene="ACHE"
/note="G00-118-746"
/number=3
/evidence=experimental
1871
/gene="ACHE"
/note="Pro446/Pro"
/replacement="t"
1994.2988
/gene="ACHE"
/note="G00-118-746"
/number=3
/evidence=experimental
2309
/gene="ACHE"
/note="g or a; within Alu repeat; g more common;
polymorphism at 2309"
/replacement="a"
2989.3158
/gene="ACHE"
/note="G00-118-746"
/number=4
/evidence=experimental
```

```

Intron 3159.3238
/gene="ACHE"
/note="G00-118-746"
/number=4
/evidence=experimental
3239.3991
/gene="ACHE"
/note="Includes retained Intron 5; G00-118-746"
/number=5
/evidence=experimental
3290
/gene="ACHE"
/note="Pro 561 more common; C/G polymorphism at 3290;
Pro561(CCG) or Arg 561(CGG)"
/replacement="g"
3992.>4185
/gene="ACHE"
/note="G00-118-746"
/number=6
/evidence=experimental

exon
BASE COUNT 792 a 1320 c 1236 g 837 t
ORIGIN
Query Match 62.0%; Score 1070; DB 9; Length 4185;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagccccgcagtgctcgtgcacacagcctccctgcttccccactctcctc 60
Db 95 ATGAGGCCCCCGAGTGTCTGTGCACACAGCCTTCCCTGGCTTCCCACTCCTCTC 154

QY 61 ctccctgctcctcgtggtgagagagtggggtgagggccggagagatcagagtctg 120
Db 155 CTCCTGTGGCTGTGGGTGAGAGAGTGGGGCTGAGGGCCGGAGAGATCAGAGTGTG 214

QY 121 gtgacgtgctggtggggccgctgcggggcatcgcctgaagacccccggggccctgtc 180
Db 215 GTGACGCTGCGTGGGGGGCCGCTGCGGGGCAATTCGCTGAAGACCCCCGGGGCCCTGTC 274

QY 181 tctgcttccctgggcatcccttgcgagaccaccatggaaccccgctccttctgcca 240
Db 275 TCTGCTTCTGTGGGCATCCCTTTGCGGAGCCACCCATGGAGCCCCCGTCTTCTGCCA 334

QY 241 ccgagcccaagcagccttggtcaggggtgtagacgtacaaccttcagagtgctgc 300
Db 335 CCGAGCCCAAGCAGCCTTGTGTCAGGGGTGTAGACGCTACAACTTCAGAGTGTCTGC 394

QY 301 taccatatgtgacaccctataccaggttttagggcaccagatgtgaaacccaac 360
Db 395 TACCAATATGTGACACCCCTATACCCAGGTTTGAAGGGCACCGAGATGTGAACCCCAAC 454

QY 361 cgtgagctgagcagagactgcctgtacctcaacgtgtgacacacataccccggcctaca 420
Db 455 CGTGAGCTGAGCGAGAGACTGCCTGTACCTCAACGTGTGACACACCATACCCCGCCTACA 514

QY 421 tccccaccctgtcctcgtcgtgatctatggtggtgtgcttctacagtggggcctctcc 480
Db 515 TCCCCACCCCTGTCTCGTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCTCC 574

QY 481 ttgacgtgtacagatggcgccttctgtgacagggccgagagagactgtgtgttcacatg 540
Db 575 TTGACGCTGTACGATGGCGCCTTCTGTGTACAGGCCGAGAGAGACTGTGCTGCATG 634

QY 541 aactaccggtggtgagccttggcttccctgcccctgcggggagccgagagagcccgggc 600
Db 635 AACTACCGGTGTGGAGCCTTTGGCTTCTGTGCCCTGCCGGGAGCCGAGAGCCCGGGC 694

QY 601 aatgtggtctcctggtacagaggtcggcctgcagtggtggtgcaggaacgtgacagcc 660
Db 695 AATGTGGGTCTCTGTGATCAGAGGCTGGCCCTGCAGTGGTGCAGAGAACTGGCAGCC 754

QY 661 ttcgggggtgaccccgacatcagtgaagctgtttggggagagcgcgggagccgcctcggtg 720
```

Db	755	TTCCGGGGTACCCGACATCACTGACGCTGTTGGGGAGACGCGGAGCCGCTCGGTG	814
QY	721	ggcatgcacctgtgtcccccagccgggctgttccacagggccgtgtcagagc	780
Db	815	GGCATGCACCTGCTGTCCCCGCCACGCGGGGCTGTCCACAGGGCCGTGCTCAGAGC	874
QY	781	ggtgcccccaatggaacctggygccaagtytggtatgagagccctgcgaagggccag	840
Db	875	GGTGCCCCCAATGGACCTGGGGCCACGGTGGGCATGGGAGAGGCCCGTCGCAGGGCCACG	934
QY	841	cagctggcccaaccttgyggtgtgtcctccagggcggaatgttggaaatgacacagagctg	900
Db	935	CAGCTGGCCCCACCTGTGTGGGCTGTCTCCAGGGCGGCACTGTGGGAATGACACAGAGCTG	994
QY	901	gtagcctgccttcggaacgaccagcgcaagtcctgtgtgaaccaagaatgcaacgtctg	960
Db	995	GTAGCCTGCCTTCGGACACGACCAGCGCAGGTCCTGTGTGAACCAACGAATGGCACGTGCTG	1054
QY	961	cctcaagaagcgctcttcgcggttctccttcgtgcctgtgtagatggaactcctcagt	1020
Db	1055	CCTCAAGAAAGCGCTTCCTCCGGTCTCTCTGTCGTGTGTGTAGATGAGACTTCCTCAGT	1114
QY	1021	gacaccccaagagccctcatcaacgcgagagactccaagccctgcaggt	1070
Db	1115	GACACCCACAGAGGCCCTCATCAACGCGGAGACTTCCACGGCCCTGCAGGT	1164

RESULT	10
AC011895/c	
LOCUS	172358 bp DNA linear PRI 30-SEP-2000
DEFINITION	Homo sapiens BAC clone RP11-126L15 from 7q22, complete sequence.
ACCESSION	AC011895
VERSION	AC011895.4 GI:9454624
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 172358)
JOURNAL	Sulston, J.E. and Waterston, R.
MEDLINE	Toward a complete human genome sequence
REFERENCE	Genome Res. 8 (11), 1097-1108 (1998)
AUTHORS	99063792
TITLE	2 (bases 1 to 172358)
JOURNAL	Sun, H., Nguyen, C., Drone, K., Laplant, Y. and Reltz, L.
REFERENCE	The sequence of Homo sapiens BAC clone RP11-126L15
AUTHORS	Unpublished
TITLE	3 (bases 1 to 172358)
JOURNAL	Waterston, R.H.
REFERENCE	Direct Submission
AUTHORS	Submitted (15-OCT-1999) Genome Sequencing Center, Washington
TITLE	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL	MO 63108, USA
REFERENCE	4 (bases 1 to 172358)
AUTHORS	Waterston, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (26-JUL-2000) Genome Sequencing Center, Washington
REFERENCE	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
AUTHORS	MO 63108, USA
TITLE	5 (bases 1 to 172358)
JOURNAL	Waterston, R.H.
REFERENCE	Direct Submission
AUTHORS	Submitted (10-SEP-2000) Genome Sequencing Center, Washington
TITLE	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL	MO 63108, USA
REFERENCE	6 (bases 1 to 172358)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (30-SEP-2000) Department of Genetics, Washington
REFERENCE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
AUTHORS	On Jul 26, 2000 this sequence version replaced gi:7630854.
TITLE	----- Genome Center
JOURNAL	
COMMENT	

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@watson.wustl.edu
----- Summary Statistics -----
Center project name: H_NH0126L15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-336D7. Actual start of this clone is at base position 1 of RP11-126L15; actual end is at base position 172358 of RP11-126L15.

The clone RP11-126L15 contains a transposon that inserts at base position 30153, which has been omitted from the submitted sequence. There are polymorphic base pair differences in the overlap between the clone RP11-126L15 and RP11-336D7.

```
FEATURES
source
    Location/Qualifiers
      1. .172358
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="7"
        /map="7q22"
        /clone="RP11-126L15"
        /clone_1lb="RPCI-11"
      1. .249
        /rpt_family="Alu"
          935..1075
            /rpt_family="Alu"
              1076..1390
                /rpt_family="Alu"
                  1391..1553
                    /rpt_family="Alu"
                      2087..2371
                        /rpt_family="Alu"
                          3027..3324
                            /rpt_family="Alu"
                              3357..3491
                                /rpt_family="Alu"
```

```
repeat_region      3492..3795      /rpt_family="Alu"
repeat_region      3799..4102      /rpt_family="Alu"
repeat_region      4550..4853      /rpt_family="Alu"
repeat_region      4879..5089      /rpt_family="Alu"
repeat_region      5107..5403      /rpt_family="Alu"
repeat_region      5495..5782      /rpt_family="Alu"
repeat_region      6357..6656      /rpt_family="Alu"
repeat_region      6661..6848      /rpt_family="Alu"
repeat_region      6849..7025      /rpt_family="Alu"
repeat_region      7770..8089      /rpt_family="Alu"
repeat_region      8110..8419      /rpt_family="Alu"
misc_feature       8185..8623      /rpt_family="Alu"
repeat_region      9602..9902      /note="similar to
repeat_region      10324..10630      /rpt_family="Alu"
repeat_region      11217..11526      /rpt_family="Alu"
repeat_region      11667..11976      /rpt_family="Alu"
repeat_region      12357..12655      /rpt_family="Alu"
repeat_region      12657..12781      /rpt_family="Alu"
repeat_region      12783..12955      /rpt_family="Alu"
repeat_region      12956..13236      /rpt_family="Alu"
repeat_region      13237..13367      /rpt_family="Alu"
repeat_region      13403..13444      /rpt_family="Alu"
repeat_region      13445..13547      /rpt_family="Alu"
repeat_region      13561..13864      /rpt_family="Alu"
repeat_region      14215..14326      /rpt_family="Alu"
repeat_region      14327..14591      /rpt_family="Alu"
repeat_region      14592..14674      /rpt_family="L1"
repeat_region      14675..14966      /rpt_family="Alu"
repeat_region      14967..15001      /rpt_family="L1"
repeat_region      15010..15337      /rpt_family="Alu"
repeat_region      16240..16509      /rpt_family="Alu"
repeat_region      16522..16656      /rpt_family="Alu"
repeat_region      16687..16978      /rpt_family="Alu"
repeat_region      17053..17344      /rpt_family="Alu"
repeat_region      18167..18444      /rpt_family="Alu"
repeat_region      18848..19163      /rpt_family="Alu"
misc_feature       19101..19232
```

EST AI539956 (NID:g4457329)"

```
/note="similar to EST BE160273 (NID:g8622994)"
repeat_region      19172..19471      /rpt_family="Alu"
repeat_region      19482..19793      /rpt_family="Alu"
repeat_region      19889..20004      /rpt_family="Alu"
repeat_region      20155..20444      /rpt_family="Alu"
repeat_region      20455..20725      /rpt_family="Alu"
```

Query Match 62.0%; Score 1070; DB 9; Length 172358;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 atgagcccccagtgctctgtcacagcctccctgtctcccactctctc 60
Db 136364 ATGAGGCCCCGACAGTGTCTGTGCACAGCCTTCCCTGGCTTCCCACTCTCTCCTC 136305

QY      61 ctctctgtctctctgtgtgagagtggtgagggccggagagatgcagagctgtc 120
Db 136304 CTCCTCTGGCTTCCTGGGTGAGAGAGTGGGGCTGAGGGCCGGAGAGATGCAGAGCTGCTG 136245

QY      121 gtacggtgcgtggggccggtgcggggcattcgctgaagaaccccggggcccctgtc 180
Db 136244 GTGACGGTGCCTGGGGGGCCGGCTGCGGGCATTCGCTGAAGACCCCGGGGGCCCTGTGC 136185

QY      181 tctgcttctctggcatcccttgcggagccaaccatggagaccgcctctctgtcca 240
Db 136184 TCTGCTTCTCTGGGCATCCCTTTGGCGGACCACCCATGGAGCCCGCTTCTTGCCA 136125

QY      241 ccgagcccaagcagccttggtcaggggtgtgagcgtacaacccctccagagtgtctgc 300
Db 136124 CCGAGCCCAAGCAGCCTTGTGTCAGGGGTGTGACGCTACAACCTTCCAGAGTGTCTGC 136065

QY      301 taccaatatgtggaacacctatacccaagttttgagggcacccgagatgtgaaccccaac 360
Db 136064 TACCAATATGTGGAACACCTTATACCCAGGTTTGAAGGGCACCCGAGATGTGAACCCCAAC 136005

QY      361 cgtgagctgagcagagactgcctgtactcaacgltgtgacaaccatacccccggcctaca 420
Db 136004 CGTGAGCTGAGCGAGGACTGCTGTACTCAACGTGTGACACCAATACCCCGGCTTACA 135945

QY      421 tcccccaacctgtcctcgtctgatactatgggggtgtgctctacagtgggcctctcc 480
Db 135944 TCCCCCACCCCTGCTCTGCTGTGATATGGGGGTGGCTTACAGTGGGGCCTCTCTCC 135885

QY      481 ttgacgtgtacgaltgcccgtctctgtacagggccgagagactgtgtgtccatg 540
Db 135884 TTGACGTGTACGATGGCCGCTTGTGATACAGGCCGAGAGAGACTGTGTGTCCATG 135825

QY      541 aactaccggtggagaccttgtgcttccctggccctgcggggagccgagagcccgggc 600
Db 135824 AACTACCGGTTGGGAGCCTTTGGCTTCTGGCCCTGCCGGGAGCGAGAGCCCGGGGC 135765

QY      601 aatgtggtctcctgtgatacagagctggccctgcagtggtgtcagagagaacgtgcaacc 660
Db 135764 AATGTGGTCTCTCGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAACTGGCACACC 135705

QY      661 ttcggggtgacccgacatacagtgaagctgtttgggagagcgcggaagccgcctcggtg 720
Db 135704 TTCGGGGGTGACCCGACATCAGTACGCTGTTTGGGAGAGCGCGGAGCCGCTCGGTG 135645

QY      721 ggcatacactgtcttccccgcccagccggggcctgttccacagggccgltgtcagagc 780
Db 135644 GGCATGCACTGTCTCCCCGCGCCAGCGGGGCGCTGTCCACAGGGCCGTGCTGCAGAGC 135585

QY      781 ggtgccccaatgagaccctgggccaagtggtgcatgtggagagggccgctgcagggccacg 840
Db 135584 GGTGCCCCAATGGAACCTTGGGCCACGGTGGGCAATGGAGAGGGCCCGTCCGAGGGCCACG 135525

QY      841 cagctggcccaacctgtgtggtctctccaggcggaactgtgtggaatatgacacagagctg 900
```


|||||
Db 135524 CAGCTGGCCACCTGTGTGGCTGTCTCCAGGCGGACACTGTGGGAATGACACAGAGCTG 135465
OY 901 gtagcctgccttcgacacgacgacgagtcctgtgtgaaccaaatgacagctgtg 960
Db 135464 GTAGCCTGCTTCGGACACGACGACGAGGTCCGTGGTGAACCAAGATGGACAGTGTGCTG 135405
OY 961 cctcaagaagaagcgtcttcgcgtcttccttcgtgcctgtgtgtagatgagacttcctcagt 1020
Db 135404 CCTCAAGAAAGCGTCTTCGGTCTCTCCTTCGTGCTGTGGTAGATGAGACTTCCTCAGT 135345
OY 1021 gacacccagagagccctcatcaacgcgggagacttcacagcgctgcaggt 1070
Db 135344 GACACCCACAGAGGCCCTCATCAACGGGGAGACTTCCACGGCCTGCAGGT 135295

RESULT 11
AC084057 194681 bp DNA linear HTG 12-OCT-2000
LOCUS Homo sapiens chromosome 7 clone CTC-786K12, WORKING DRAFT SEQUENCE,
DEFINITION 36 unordered pieces.
AC084057 AC084057.1 GI:10799406
VERSION AC084057.1 GI:10799406
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 194681)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 7
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 194681)
REFERENCE DOE Joint Genome Institute.
AUTHORS Direct Submission
TITLE Submitted (12-OCT-2000) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 2351286
Center clone name: CTC-786K12

Summary Statistics
Consensus quality: 162650 bases at least Q40
Consensus quality: 173899 bases at least Q30
Consensus quality: 176608 bases at least Q20
Estimated insert size: 200000; pulse field gel estimation
Estimated coverage: 7.95 in Q20 bases; pulse field gel estimation
Quality coverage: 8.32 in Q20 bases; sum-of-contigs estimation
Quality coverage: 8.32 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1324: contig of 1324 bp in length
* 1325 1424: gap of unknown length
* 1425 2586: contig of 1162 bp in length
* 2587 2686: gap of unknown length
* 2687 4056: contig of 1370 bp in length
* 4057 4156: gap of unknown length
* 4157 5197: contig of 1041 bp in length
* 5198 5297: gap of unknown length
* 5298 6413: contig of 1116 bp in length
* 6414 6513: gap of unknown length
* 6514 7675: contig of 1162 bp in length

* 7676 7775: gap of unknown length
* 7776 8843: contig of 1068 bp in length
* 8844 8943: gap of unknown length
* 8944 9998: contig of 1055 bp in length
* 9999 10098: gap of unknown length
* 10099 11598: contig of 1500 bp in length
* 11599 11699: gap of unknown length
* 11699 12956: contig of 1258 bp in length
* 12957 13056: gap of unknown length
* 13057 14168: contig of 1112 bp in length
* 14169 14268: gap of unknown length
* 14269 15287: contig of 1019 bp in length
* 15288 15387: gap of unknown length
* 15388 16936: contig of 1549 bp in length
* 16937 17036: gap of unknown length
* 17037 18086: contig of 1050 bp in length
* 18087 18186: gap of unknown length
* 18187 19427: contig of 1241 bp in length
* 19428 19527: gap of unknown length
* 19528 20570: contig of 1043 bp in length
* 20571 20670: gap of unknown length
* 20671 22193: contig of 1523 bp in length
* 22194 22293: gap of unknown length
* 22294 24473: contig of 2180 bp in length
* 24474 24573: gap of unknown length
* 24574 25995: contig of 1422 bp in length
* 25996 26095: gap of unknown length
* 26096 27843: contig of 1748 bp in length
* 27844 27943: gap of unknown length
* 27944 30120: contig of 2177 bp in length
* 30121 30220: gap of unknown length
* 30221 31961: contig of 1741 bp in length
* 31962 32061: gap of unknown length
* 32062 35719: contig of 3658 bp in length
* 35720 35819: gap of unknown length
* 35820 38071: contig of 2252 bp in length
* 38072 38171: gap of unknown length
* 38172 41171: contig of 3000 bp in length
* 41172 41271: gap of unknown length
* 41272 47552: contig of 6281 bp in length
* 47553 47652: gap of unknown length
* 47653 52017: contig of 4365 bp in length
* 52018 52117: gap of unknown length
* 52118 55609: contig of 3492 bp in length
* 55610 55709: gap of unknown length
* 55710 60279: contig of 4570 bp in length
* 60280 60379: gap of unknown length
* 60380 66246: contig of 5867 bp in length
* 66247 66346: gap of unknown length
* 66347 78184: contig of 11838 bp in length
* 78185 78284: gap of unknown length
* 78285 96184: contig of 17900 bp in length
* 96185 96284: gap of unknown length
* 96285 115729: contig of 19445 bp in length
* 115730 115829: gap of unknown length
* 115830 139049: contig of 23220 bp in length
* 139050 139149: gap of unknown length
* 139150 162002: contig of 22753 bp in length
* 161903 162002: gap of unknown length
* 162003 194681: contig of 32679 bp in length.
location/Qualifiers
1. 194681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/clone="CTC-786K12"
/clone_11b="CalTech human BAC library C"
BASE COUNT 45257 a 49777 c 50799 g 45331 t 3517 others
ORIGIN
Query Match 62.0%; Score 1070; DB 2; length 194681;
Best Local Similarity 100.0%; Pred. No. 0;

/note="Cpg Island score = 0.87, GC = 67.80%, CpGs = 13"
/note="Region: Cpg Island"
/evidence=not_experimental
Intron 2105..2434
/note="Alignment: EST, | Intron"
/evidence=not_experimental
Intron 2497..2596
/note="Alignment: EST, | Intron"
/evidence=not_experimental
exon 2597..2675
/note="Xpound exon prediction, score = 98% (0%)"
/evidence=not_experimental
exon 2597..2675
/note="MZEf, score = 98.9%"
/evidence=not_experimental
Intron 2676..2980
/note="Alignment: EST, | Intron"
/evidence=not_experimental
Intron 3170..3528
/note="Alignment: EST, | Intron"
/evidence=not_experimental
exon 3529..3585
/note="MZEf, score = 100%"
/evidence=not_experimental
misc_feature 3539..3764
/note="Cpg Island score = 0.82, GC = 62.80%, CpGs = 18"
/note="Region: Cpg Island"
/evidence=not_experimental
exon 3548..3702
/note="Xpound exon prediction, score = 98% (0%)"
/evidence=not_experimental
exon 3637..3702
/note="Alignment: CDS, |CG415001|87.9|46|111"
/evidence=not_experimental
Intron 3703..3780
/note="Alignment: CDS, | Intron"
/evidence=not_experimental
exon 3781..3934
/note="MZEf, score = 97.3%"
/evidence=not_experimental
exon 3781..3934
/note="Xpound exon prediction, score = 99% (0%)"
/evidence=not_experimental
exon 3781..3927
/note="Alignment: CDS, |CG415001|85.7|112|258"
/evidence=not_experimental
misc_feature 3812..3960
/note="BLASTN (EMBL), 110..258 of emb|U41500|CG41500
Cricketulus griseus arsenite-resistance protein (ars2)
mRNA,"
/note="Region: Data base match"
/evidence=not_experimental
exon 4054..4281
/note="MZEf, score = 84.8%"
/evidence=not_experimental
exon 4054..4247
/note="Xpound exon prediction, score = 98% (0%)"
/evidence=not_experimental
exon 4191..4220
/note="BLASTX (GENPEPT), 58..87 of gb|U41500|CGU41500_1
Cricketulus griseus arsenite-resistance protein (ars2)"
/evidence=not_experimental
Intron 4248..4436
/note="Alignment: CDS, | Intron"
/evidence=not_experimental
exon 4437..4593
/note="MZEf, score = 97%"
/evidence=not_experimental
exon 4437..4593
/note="Xpound exon prediction, score = 89% (0%)"
/evidence=not_experimental
exon 4437..4593
/note="Alignment: CDS, |CG415001|86.6|457|613"

misc_feature
4466..4626
/note="BLASTN (EMBL), 453..613 of emb|U41500|CG41500
Cricketulus griseus arsenite-resistance protein (ars2)
mRNA,"
/note="Region: Data base match"
/evidence=not_experimental
exon 4467..4519
/note="BLASTX (GENPEPT), 87..139 of gb|U41500|CGU41500_1
Cricketulus griseus arsenite-resistance protein (ars2)"
/evidence=not_experimental
Intron 4594..4788
/note="Alignment: EST, | Intron"
/evidence=not_experimental
exon 4809..4840
/note="BLASTX (GENPEPT), 140..171 of gb|U41500|CGU41500_1
Cricketulus griseus arsenite-resistance protein (ars2)"
/evidence=not_experimental
Intron 4879..4991
/note="Alignment: CDS, | Intron"
/evidence=not_experimental
exon 4992..5118
/note="Alignment: CDS,
|CG415001|88.3|717|844|CG415002|88.6|1144"
/evidence=not_experimental
5045..5152
/note="BLASTN (EMBL), 738..845 of emb|U41500|CG41500
Cricketulus griseus arsenite-resistance protein (ars2)
mRNA,"
/note="Region: Data base match"
/evidence=not_experimental
Intron 5119..5208
/note="Alignment: CDS, | Intron"
/evidence=not_experimental
5209..5264
/note="Alignment: CDS, |CG415001|89.3|845|900"
/evidence=not_experimental
misc_feature
complement(5238..5431)
/note="BLASTN (EMBL), 1..194 of emb|G30851|HS1140038 human
chromosome 7 STS SWS3840: single read."
/note="Region: Data base match"
/evidence=not_experimental
5757..14176
/note="GC score = 489.57 (8420bp)"
/note="Region: GC content"
/evidence=not_experimental
5844..7208
/note="Cpg Island score = 0.79, GC = 69.50%, CpGs = 126"
/note="Region: Cpg Island"
/evidence=not_experimental
6019..6471
/note="Cpg core score = 8.90"
/note="Region: Cpg Island"
/evidence=not_experimental
6454..6508
/note="homology = 89.10%, score = 38, counts = 11"
/rpt_type=tandem
/rpt_unit=ccctcc
repeat_region
misc_feature
Query Match 45.3%; Score 781; DB 9; Length 34921;
Best local Similarity 100.0%; Pred. No. 0;
Matches 781; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagcccccgagctgtctgtctgacacagccttcctggtctcccaactcctctc 60
|||||
Db 10809 ATGAGGCCCCGAGTGTCTGTCACACGCCCTTCCCTGGCTTCCCACTCTCTCCTC 10750
QY 61 ctctctgtctcctgtgtggaagagtggggtgagggcgaggagatgcagagctgtc 120
|||||
Db 10749 CTCCTCTGCTCCTGCTGAGGAGAGTGGGGCTGAGGGCCGGAGAGATGCAGAGCTGCTG 10690
QY 121 gtgacggtgcgtgtgggacggctgcggggcatctgcctgaagacccccgggggcctgtc 180
|||||

Db	Accession	Sequence	Length
Db	10689	GTGACGGGTGCGTGGGGGGCGGCTGCGGGCATTCGCGCTGAAGACCCCCGGGGCCCTGTC	10630
QY	181	tcgtcttctctggtgcatccccccttgcgagaccaccatggaaccccgctcgtcttcgca	240
Db	10629	TCCTGCTTTCCTGGGCATCCCCCTTTGCGGAGCCACCACCATGGGACCCCGCTTCTGCCA	10570
QY	241	ccggagccccaagcagcccttggtcagggtgtgtagacgtacaaccttccagagtgtctgc	300
Db	10569	CCGGAGCCCAAGCAGCCCTGTGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGC	10510
QY	301	taccaatatgttgacaacccctatacccaagtctttagaggcaccgagatgtgaaacccaac	360
Db	10509	TACCAATATGTGGACACCCCTATACCCAGGTTTGAAGGGCACCGAGATGTGGAACCCCAAC	10450
QY	361	cgtagcgtgagcgaaggaactgctgttaacctcaacgtgtgacaccataacccccgacctaca	420
Db	10449	CGTAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGACACCATACCCCGGCGCTACA	10390
QY	421	tcceccacccctgtcctcgtctgatactataggggtgtgctctacagltgaggcctctcc	480
Db	10389	TCCCCACCCCTGTCTCTGTTGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCTCC	10330
QY	481	ttggaacgtgtacgatgagccgtctcttggtacagagccgagaggaactgtgtgtccatg	540
Db	10329	TTGGACGTGTACGATGAGCGCGCTTCTGTGTACAGAGCCGAGAGACTGTGCTGTCTCATG	10270
QY	541	aactaccgggtgtggagacgttggcttccctgtgcccctgcgggggagccgagagcccccggc	600
Db	10269	AACTACCGGGGTGGAGACCTTTGGCTTCTCTGGCCCTGCCGGGAGACCCGAGAGGCCCGGGC	10210
QY	601	aatgtgggtctcctgtgatcagaaggtgcctgcctcagttggtgagagaggaactgtgcagcc	660
Db	10209	AATGTGGGTCTCTCTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC	10150
QY	661	ttcgggggtgaccccgacatcagtgaagctgttttgaggagagcgaggagccgctcggtg	720
Db	10149	TTGGGGGTGACCCGACATCAGTACGCTGTTTGGGAGAGCGCGGAGCGCGCTCGGTG	10090
QY	721	ggcatgcacctgtgtcccccacagccgagcggtgttccacaagggccgtgtcagagc	780
Db	10089	GGCATGCACCTGTCTGTCCCGCCAGCGGGGCGCTGTTCACAGGGCGCGTGTGACAGC	10030
QY	781	g 781	
Db	10029	G 10029	

REFERENCE	3 (bases 1 to 779)
AUTHORS	Bartels,C.F., Zelnickl,T. and Lockridge,O.
TITLE	Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene accounts for YT blood group polymorphism
JOURNAL	Am. J. Hum. Genet. 52 (5), 928-936 (1993)
MEDLINE	93256075
COMMENT	On Jun 8, 1995 this sequence version replaced gi:348160. **Reference [1] reports bases 99-583; Reference [J. Biol. Chem. 266, 23083-23090 (1991)] reports bases 89-98 and 584-592; Reference [3] reports bases 593-781**.
FEATURES	Location/Qualifiers source 1..779 /organism="Homo sapiens" /db_xref="taxon:9606" /map="7q22" /tissue_type="blood" order(L22559.1:1225..>1355,<1..97) /note="G00-118-746" /number=2 exon 98..582 /gene="ACHE" /note="G00-118-746" /number=3 variation 460 /gene="ACHE" /note="C to T (Pro/Pro) polymorphism" /replace="t"
BASE COUNT	159 a 214 c 259 g 147 t
ORIGIN	,
Query Match	28.3%; Score 489; DB 9; Length 779;
Best Local Similarity	100.0%; Pred. No. 7e-257;
Matches	489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1066 cagtgctgctggtggtgtgtgaagtatgaggctcgattttcgtgttacggggcccca 1125 Db 95 CAGGTGCTGGTGGGTGTGTGAAGATGAGGGCTCGATTTCGTGTACGGGGCCCA 154
OY	1126 ggcttcagcaagaacaagatctctcatcagccgggcgaagttccttggccgggtgcg 1185 Db 155 GGCTTCAGCAAAGACAACGACTCTCATCAGCCGGGCCGAGTTCTTGCGGGGTGCG 214
OY	1186 gtccgggttcccccaagtaagtaccttgcagccgaagcttgttccctgcattacacagac 1245 Db 215 GTCCGGGTTC CCCAGTAAGTGCCTGGCAGCCGAGGCTGTGTCTGCATTACACAGAC 274
OY	1246 ttgctgcattcccagagaccggcacgccctgagggaagccctagagcatgtgttggcgac 1305 Db 275 TTGCTGCATCCCCGAGACC CGCACGCCCTGAGGAGGGCCCTGAGCGATGTGTGGGCCAC 334
OY	1306 caaatgtcgtgtgccccgtyggccacagctgtgctgggcgaactgttcccagggtgccgg 1365 Db 335 CACAATGTCTGTGCCCCCGTG GCCCAGCTGGCTGGGCGACTGGCTGCCAAGGTTGCCGG 394
OY	1366 gtctaagcctaagctcttgaacaacccgtgtctcaacgctctccttggccctgtgatggg 1425 Db 395 GTCTACGCGCTACGCTTTTGAACACC GTGCTTCCACGCTCTCTGGCCCCCTGTGATGGGG 454
OY	1426 gtgccccacggtacagagatcgaattcatcttttgatcccccttgagaccctctgaaac 1485 Db 455 GTGCCCCACGGCTACGAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCGAAAC 514
OY	1486 tacacggcagagagaataatcttcgcccacagcgaactgatgtcgaatacttggccaactttgc 1545 Db 515 TACACGGCAGAGAGAATAATCTTCGCCCAAGCAGCTGATGCGATACTGGCCAAC TTGCC 574
OY	1546 cgcacaggg 1554 Db 575 CGCACAGGG 583

RESULT 14
LOCUS HUMACHE03 510 bp DNA linear PRI 14-JUN-1995
DEFINITION Human acetylcholinesterase (ACHE) gene, exons 4 and 5.
ACCESSION L22561
VERSION L22561.1 GI:862301
KEYWORDS acetylcholinesterase.
SEGMENT 3 of 4
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Soreq,H.E., Ben-Aziz,R., Prody,C.A., Seldman,S., Gnatt,A.,
Neville,L., Lleman-Hurwitz,J., Lev-Lehman,E., Ginzberg,D.,
Lapidot-Lifson,Y. and Zakut,H.
TITLE Molecular cloning and construction of the coding region for human
acetylcholinesterase reveals a G + C-rich attenuating structure
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9688-9692 (1990)
MEDLINE 91088577
REFERENCE 2 (sites)
AUTHORS Li,Y., Camp,S., Rachinsky,T.L., Getman,D. and Taylor,P.
TITLE Gene structure of mammalian acetylcholinesterase. Alternative exons
dictate tissue-specific expression
JOURNAL J. Biol. Chem. 266 (34), 23083-23090 (1991)
MEDLINE 92078174
REFERENCE 3 (bases 1 to 510)
AUTHORS Bartels,C.F., Zelinski,T. and Lockridge,O.
TITLE Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene
accounts for YF blood group polymorphism
JOURNAL Am. J. Hum. Genet. 52 (5), 928-936 (1993)
MEDLINE 93256075
COMMENT On Jun 15, 1995 this sequence version replaced gi:853591.
**Reference [1] reports bases 106-275;
Reference [J. Biol. Chem. 266, 23083-23090 (1991)] reports bases
276-507;
Reference [3] reports bases 1-105**.
FEATURES
SOURCE Location/Qualifiers
1..510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7q22"
/tissue_type="blood"
order(L22560..1:583..>779,<1..106)
/note="G00-118-746"
/number=3
exon 107..276
/gene="ACHE"
/note="G00-118-746"
/number=4
Intron 277..356
/gene="ACHE"
/note="G00-118-746"
/number=4
exon 357..>510
/gene="ACHE"
/note="glycolipid anchored form; G00-118-746"
/number=5
variation 408
/gene="ACHE"
/note="C to G (Pro561 to Arg) polymorphism"
/replace="g"
BASE COUNT 65 a 204 c 150 g 91 t
ORIGIN
Query Match 9.9%; Score 170; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.4e-81;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 gggatcccaatgagcccgagaccccaagggcccaatgcccccggtacacgcgggg 1612
|||||
Db 106 GGATGCCAATGAGCCCGAGACCCCAAGGCCCCACAAATGGCCCCGTACACGGCGGGG 165

QY 1613 ctacagcagctagttagctggaacctgsgggccgctggaagtgcgsggggctgcgcgcc 1672
|||||
Db 166 CTCAGCAGTAGCTAGTCTGACACCTGGCGCGCTGAGGTGCGGGGGGCTGCGGCC 225
QY 1673 aggcctgcgctcttggaaccgcttcctcccaattgctcagcgccacc 1722
|||||
Db 226 AGGCTGCGCCTTCTGGAACCGCTTCCTCCCAATTGCTCAGCGCCACC 275
RESULT 15
LOCUS AF312032 117962 bp DNA linear PRI 20-MAR-2001
DEFINITION Homo sapiens acetylcholinesterase (ACHE) gene, partial cds; ASR2
(ASR2) gene, complete cds, alternatively spliced; thyroid receptor
interacting protein 6 (TRIP6) gene, complete cds; cation-chloride
cotransporter (CIP1) gene, complete cds, alternatively spliced;
ephrln type-B receptor 4 precursor (EPHB4) and zonadhesin (ZAN)
genes, complete cds; and unknown gene.
ACCESSION AF312032
VERSION AF312032
KEYWORDS AF312032.1 GI:13383497
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 117962)
AUTHORS Wilson,M.D., Riemer,C., Martindale,D.W., Schnupf,P., Boright,A.P.,
Cheung,T.L., Hardy,D.M., Schwartz,S., Scherer,S.W., Tsui,L.-C.,
Miller,W. and Koop,B.F.
TITLE Comparative analysis of the gene-dense ACHE/TRF2 region on human
chromosome 7q22 with the orthologous region on mouse chromosome 5
Nucleic Acids Res. 29 (6), 1352-1365 (2001)
JOURNAL 21138439
MEDLINE 2 (bases 1 to 117962)
AUTHORS Wilson,M.D., Martindale,D.W. and Koop,B.F.
TITLE Direct Submission
COMMENT Submitted (06-OCT-2000) Biology, Centre for Environmental Health,
University of Victoria, P.O. Box 3020, Victoria, B.C. V8W 3N5,
Canada
FEATURES
SOURCE Location/Qualifiers
1..117962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q22"
/clone="cos159d9.L and H_DJ0138m12 fusion"
join(<1..165,999..1339)
/gene="ACHE"
/product="acetylcholinesterase"
<1..1339
/gene="ACHE"
join(<1..165,999..1120)
/gene="ACHE"
/codon_start=3
/product="acetylcholinesterase"
/protein_id="AAK21003.1"
/db_xref="GI:13383498"
/translation="NEPRDPKAPQWPPTAGAOQYVSLDRLPLEVRRLRAQACAFWN
RELPLKLSATDITLDEAERQWKAEFHRWSSYVHWKMGFDHYSKQDRCSDL"
1676..2594
/product="unknown"
2048..2476
/codon_start=1
/evidence="not_experimental"
/product="unknown"
/protein_id="AAK21004.1"
/db_xref="GI:13383499"
/translation="MGDKRPGFRGSRDWTGCVESICLAHFGPGGRICHYRGGVLH
GELRLYSHFAGGGPVWVGADADARSKALGCVSGTEAYLVLDPHYWGTPKSPS
ELQAAGWVGQEVSAFDPNSEFYNLCITSLSSQQQRTLD"
3460..3616,3806..3999,4119..4272,4351..4524,4884..5072,

5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15746,15990. .16352))
/gene="ASR2"
/note="alternatively spliced; missing 12 bp from exon 18"
/product="ASR2B"
complement(join(2655. .2845,2936. .3062,3176. .3277,
3460. .3616,3806. .3999,4119. .4272,4351. .4524,4884. .5072,
5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15746,15990. .16352))
/gene="ASR2"
/note="alternatively spliced"
/product="ASR2A"
complement(2655. .16352)
/gene="ASR2"
complement(join(2770. .2845,2936. .3062,3176. .3277,
3460. .3616,3806. .3999,4119. .4272,4351. .4524,4884. .5072,
5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15728))
/gene="ASR2"
/note="alternatively spliced"
/codon_start=1
/product="ASR2A"
/protein_id="AAK21005.1"
/db_xref="GI:13383500"
/translation="MGSDSDDEYDRRRRDKFRERSDYDRSRERDRRRGDDWNDREWD
RGRERSRGEYRDYDRNRERESPPRHLSPPQKMRMRDDEHSSDPYHSGYEMPYAG
GGGPTYGPPQPMGHPDVHIMQHVLPIQARLGSIAEIDLGVPPVMKTFKEFLSLD
DSVDETAVKRYNDYKLDFFRQOMDFFLAHKDEEWFRRKYHPDEVGKRRQEARGALO
NRLRFLSLMETGWFNDLLDIDKADAIKMLDAVIKMEGTENDLRILEQEEEEEQ
AGKPEPSKKEEGRAGAGLDEGRKTNDKEKKEDGQAENDSNDKTKKSEGDDK
EEKKDESEKAKKSSKRNKRKSGDDSFDEGSVSESESESGQAEEKEEAEEALKE
KEKPEEWEKPKDAAGLECKRPLHKTCSLFMRNIAPNISRAEIIISLCRYPGFMRV
ALSEPQPERFRFRGWTFDRSNIKEICWNLQNIIRLRECELSPGVNRDLTRVRNIN
GITQHKQIVRNDIKLAAKLIHLTDRTQMWASEPPTPLPSLSONPILKNTIDYLI
EEVSAEEELLGSSGAPPEEPKEGNPAPINVERDEKLIVLKLILYLRIVHSLDY
YNTCEYPNEDEMPNRCGIHYGPMPPNRISHGEVLEWQTFEEKLTPLSVRESLSE
EEAQKGRKDPQOEVEKFTVSTQELGDKWLCPLSGKFRGPEFVRKHIFNKHAEKI
EEVKKEVAFENNFLTDAKRPALEIKPAOPGPAQILPPGLTGLPYPHQTPOGLMPY
GQPPRPILGYGAGAVRPAPVPGPPYPHAPYGAGRGNYDAFRGQGGYPGKPRNRMVRG
DPRAIYERDLADDDVDF"
complement(join(2770. .2845,2936. .3062,3176. .3265,
3460. .3616,3806. .3999,4119. .4272,4351. .4524,4884. .5072,
5378. .5456,5557. .5618,5949. .6106,6279. .6396,6481. .6579,
6767. .6951,7080. .7249,9078. .9266,9514. .9660,9906. .10034,
15607. .15728))
/gene="ASR2"
/note="alternatively spliced; missing 12 bp from exon 18"
/codon_start=1
/product="ASR2B"
/protein_id="AAK21006.1"
/db_xref="GI:13383501"
/translation="MGSDSDDEYDRRRRDKFRERSDYDRSRERDRRRGDDWNDREWD
RGRERSRGEYRDYDRNRERESPPRHLSPPQKMRMRDDEHSSDPYHSGYEMPYAG
GGGPTYGPPQPMGHPDVHIMQHVLPIQARLGSIAEIDLGVPPVMKTFKEFLSLD
DSVDETAVKRYNDYKLDFFRQOMDFFLAHKDEEWFRRKYHPDEVGKRRQEARGALO
NRLRFLSLMETGWFNDLLDIDKADAIKMLDAVIKMEGTENDLRILEQEEEEEQ
AGKPEPSKKEEGRAGAGLDEGRKTNDKEKKEDGQAENDSNDKTKKSEGDDK
EEKKDESEKAKKSSKRNKRKSGDDSFDEGSVSESESESGQAEEKEEAEEALKE
KEKPEEWEKPKDAAGLECKRPLHKTCSLFMRNIAPNISRAEIIISLCRYPGFMRV
ALSEPQPERFRFRGWTFDRSNIKEICWNLQNIIRLRECELSPGVNRDLTRVRNIN
GITQHKQIVRNDIKLAAKLIHLTDRTQMWASEPPTPLPSLSONPILKNTIDYLI
EEVSAEEELLGSSGAPPEEPKEGNPAPINVERDEKLIVLKLILYLRIVHSLDY
YNTCEYPNEDEMPNRCGIHYGPMPPNRISHGEVLEWQTFEEKLTPLSVRESLSE
EEAQKGRKDPQOEVEKFTVSTQELGDKWLCPLSGKFRGPEFVRKHIFNKHAEKI
EEVKKEVAFENNFLTDAKRPALEIKPAOPGPAQILPPGLTGLPYPHQTPOGLMPY
GQPPRPILGYGAGAVRPAPVPGPPYPHAPYGAGRGNYDAFRGQGGYPGKPRNRMVRG
DPRAIYERDLADDDVDF"

/gene="TRIP6"
/product="thyroid receptor interacting protein 6"
complement(17865. .23970)
/gene="TRIP6"
complement(join(18015. .18146,18574. .18694,19596. .19774,
20574. .20743,20854. .20947,22453. .22824,23086. .23211,
23331. .23458,23713. .23821))
/gene="TRIP6"
/codon_start=1
/product="thyroid receptor interacting protein 6"
/protein_id="AAK21007.1"
/db_xref="GI:13383502"
/translation="MSGPTWLPKQPEPARAPQGRAIPKGTGPPPAHGALQHPRV
NECPILPSEQYOAPGCPEDRGPAWGSGLQHTQGLPADRGGLRPSLDAEIDLSS
TLAELNGGQGHASRRPDRQAYEPPPPAYRTGSLKPNPASPLPASPYPGPTPASYTTA
STPAGAPFPVQVKVAQPVRCGCPRRRGASQASGPLPGRHFFLPGRGEVWGPGYSQRE
PGGAKPEAAAGVSGPAGRGREGHPQVPLSQPPEDELDRLTKLVHDMNHPSPSGEYF
GQCGGCGEDVVDGAGVALDRVHVGCFCVSTCRAQLRGOHFAVERRAYCEGCYVA
TLEKATCSQPTILDRILRAMGKAVHGPCFTCVVCHRGLDGIPFTVDATSOIHCIEDFH
RKFAKCSVCGAIIIMPEQGEYVRIVALDRSFHIGCYKEECGLLLSSEGCQCTP
LDGHILCKACSAWRIQELSATVTTDC"
complement(join(24306. .25600,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38581))
/gene="CIP1"
/note="alternatively spliced"
/product="cation-chloride cotransporter-interacting
protein 1"
complement(24306. .38581)
/gene="CIP1"
complement(join(24714. .25600,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37121))
/gene="CIP1"
/codon_start=1
/product="cation-chloride cotransporter-interacting
protein 1"
/protein_id="AAK21008.1"
/db_xref="GI:13383503"
/translation="MASESSPLHAYRLIGEVALPANGAGPGGASARKLSTFLGVV
VPYVLSMESIVFLRIGFVGHAGLLQALAMLVAFITLVLSYCAIATNGAVOGG
GAYFMISRTLGPVEGSGIGLMEFLYANVCGCAVSLGLVESVDVFGADATGPSGLRYL
PQGYGNLLYGSLLGLVGVCYTLGAGLYARASFTEFLVSGSLASVLISFVAVGPRD
IRLTPRGPNGSSLPRFGHFEGFNSSTLVKDNIGAGYAEVDTTGAVMNFASVFAVLFN
GCTGTMAGANMSGELKDPRAIPDLGITVAVAYTFVYVLLFPLSFTCDRTLQOEDYG
FFRAISLMPPLVLIGIYATALSMSLSIGASRIILHALRDLLEGGVILAPAKVVSRCG
NPMAVLYSWGLVOLVLLAGKLTLAAVTVFYLVAYAVDLSCLSEWASAPNFRPT
FSLFSWHTCLLGVASCLMMFLISPGAAGSLLMGLMALLTARGGSSWGYVSQL
LGHQVRKYLLRLDVRKDHVKFWRPQDLLLVGNPRGALPLRLANQLKGGLYLVGHVT
LGDLDSLPSDPVQPOYGAWLSLYDRAOVKAFVDLTLSPSYRQGAQHLLRISGLGMKP
NTLVLFYVDAPPODHFLLTPAPSEPADSTREGSSPALSTLPPRPAPGSPRALNPD
YVATYADALKMKNVNLARASGALPPERLSRSGSGTSQLHVDVWPLNLRPRGPGY
VDVCGFLFLQMATIILGMVPAMHSRLRIFLCIGPREAPGAEGRLRALLSOLRIRAEV
QEVVWGEAGAGAPEAAEEEGDFVNSGRDAEAEALARSANALVRAQQGRGTGGPGGP
EGDAEGPITALTLFYLPRPADPARYPRYLALLFTLTRDGLPTLLVHGVTPVTCIDL
"
complement(join(25787. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36565. .36699,36941. .37163,38485. .38541))
/gene="CIP1"
/product="cation-chloride cotransporter-interacting
protein 1 isoform b"
complement(join(25934. .25959,28491. .28637,29407. .29581,
29734. .29926,30056. .30180,31086. .31168,31276. .31433,
32157. .32268,32376. .32483,34142. .34450,35481. .35612,
36

Db 1 CCAATGAGCCCCGAGACCCCAAGGCCCCACATGGCCCCCGTACAGCGGGGGCTCAGC 60
QY 1619 agtacgttagtcttggaacctgcygcgctggaggtgcgcygggctgcgccccagacct 1678
Db 61 AGTACGTTAGTCTGGACCTGCGGGCCGCTGGAGGTGCGGGGGGCTGCGCCGCCAGGCCCT 120
QY 1679 gcgccttctggaaaccgcttcctccccaattgtctcagcgccacc 1722
Db 121 GCGCCTTCTGGAACCGCTTCCTCCCAATTGCTCAGCGCCACC 164

Search completed: August 31, 2002, 22:49:43
Job time: 26012 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 13:42:27 ; Search time 9324.87 Seconds

(without alignments)
65.081 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattcgcagccatgctagggcccccgc 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Query Score Match Length DB ID Description

1	29	100.0	29	6	AX275252	AX275252 Sequence
2	21	72.4	110000	2	LMFLCHR36_26	Continuation (27 o
3	20.2	69.7	157813	2	AC015957	AC015957 Homo sapi
4	20.2	69.7	169979	2	AC013700	AC013700 Homo sapi
5	20	69.0	165412	2	AC010811	AC010811 Homo sapi
6	20	69.0	184018	2	AC068549	AC068549 Homo sapi
7	20	69.0	185070	2	AC087279	AC087279 Homo sapi
8	19.8	68.3	110000	2	AC073604_0	AC073604 Mus muscu
9	19.8	68.3	115062	2	AC094266	AC094266 Rattus no
10	19.8	68.3	151733	2	AC079643	AC079643 Mus muscu
11	19.6	67.6	1560	8	AF321096	AF321096 Pichia pa
12	19.6	67.6	11358	6	AX175474	AX175474 Sequence
13	19.6	67.6	15873	1	AE007115	AE007115 Mycobacte
14	19.6	67.6	40121	1	MSGY414A	AD000007 Mycobacte
15	19.6	67.6	43430	1	MTCY16B7	Z81331 Mycobacteri
16	19.6	67.6	96638	9	AC092545	AC092545 Homo sapi
17	19.6	67.6	122103	2	AC084839	AC084839 Homo sapi
18	19.6	67.6	140331	9	AC007002	AC007002 Homo sapi
19	19.6	67.6	144135	9	AC097494	AC097494 Homo sapi
20	19.6	67.6	177206	2	AC098132	AC098132 Rattus no
21	19.6	67.6	178747	2	AC094025	AC094025 Rattus no
22	19.6	67.6	187113	2	AL662839	AL662839 Mus muscu
23	19.6	67.6	196855	2	AC022297	AC022297 Mus muscu
24	19.6	67.6	204135	2	AL596207	AL596207 Mus muscu
25	19.4	66.9	113891	9	AC004797	AC004797 Homo sapi
26	19.4	66.9	194509	8	AP001129	AP001129 Oryza sat
27	19.4	66.9	203563	2	AC069454	AC069454 Homo sapi
28	19.4	66.9	227518	2	AC091287	AC091287 Mus muscu
29	19.2	66.2	375	6	AX071634	AX071634 Sequence
30	19.2	66.2	657	6	AX118907	AX118907 Sequence
31	19.2	66.2	729	9	HS342957	AJ342957 Homo sapi
32	19.2	66.2	122818	9	HS1056L3	AL031727 Human DNA
33	19.2	66.2	130080	9	AP003160	AP003160 Homo sapi
34	19.2	66.2	130615	9	AC010395	AC010395 Homo sapi
35	19.2	66.2	135248	9	AC012610	AC012610 Homo sapi
36	19.2	66.2	136906	2	RN86I20	AL603805 Rattus no
37	19.2	66.2	156880	9	AC006466	AC006466 Homo sapi
38	19.2	66.2	164176	2	AC094004	AC094004 Rattus no
39	19.2	66.2	171320	2	AC091669	AC091669 Pan trogl
40	19.2	66.2	180257	2	AC007044	AC007044 Homo sapi
41	19.2	66.2	187985	2	AC091718	AC091718 Pan trogl
42	19	65.5	254	11	AU026239	AU026239 Rattus no
43	19	65.5	1518	10	AF125176	AF125176 Mus muscu
44	19	65.5	2227	9	AF242557	AF242557 Homo sapi
45	19	65.5	2441	9	AK055571	AK055571 Homo sapi

ALIGNMENTS

RESULT 1
AX275252 LOCUS AX275252 29 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 1 from Patent WO0171014.
ACCESSION AX275252
VERSION AX275252.1 GI:16547672

KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (sites)
AUTHORS Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants

JOURNAL Patent: WO 0171014-A 1 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES Location/Qualifiers
source 1..29

/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer pAcHe-Nco, derived from human AcHe gene and

misc_feature 1. .13053 /note="assembly_fragment"
misc_feature 13154. .51004 /note="assembly_fragment"
misc_feature 51105. .157813 /note="assembly_fragment"
misc_feature 51105. .157813 /note="assembly_fragment"
clone_end:T7
vector_side:left"
BASE COUNT 50360 a 31888 c 30054 g 45308 t 203 others
ORIGIN
Query Match 69.7%; Score 20.2; DB 2; Length 157813;
Best Local Similarity 88.0%; Pred. No. 92;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 tctgcagcatgctagggcccccgc 29
Db 110234 TCTGCAGCCATGCGTCGCAGCCGC 110258
RESULT 4
AC013700 169979 bp DNA linear HTG 17-MAR-2000
LOCUS Homo sapiens clone RP11-22C7, WORKING DRAFT SEQUENCE, 14 unordered
pieces.
AC013700
AC013700.4 GI:7259721
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 169979)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-22C7
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 169979)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Liew,C., Locke,K., Macdonald,P., Margulis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Titrrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 17, 2000 this sequence version replaced gi:6970506.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4099
Center clone name: 22_C_7
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

Consensus quality: 160794 bases at least Q40
Consensus quality: 165604 bases at least Q30
Consensus quality: 167325 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 168679; sum-of-contigs
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1339: contig of 1339 bp in length
1340 1439: gap of 100 bp
1440 2542: contig of 1103 bp in length
2543 2642: gap of 100 bp
2643 4104: contig of 1462 bp in length
4105 4204: gap of 100 bp
4205 5645: contig of 1441 bp in length
5646 5745: gap of 100 bp
5746 10797: contig of 5052 bp in length
10798 10897: gap of 100 bp
10898 18365: contig of 7468 bp in length
18366 18465: gap of 100 bp
18466 28722: contig of 10257 bp in length
28723 28822: gap of 100 bp
28823 40013: contig of 11191 bp in length
40014 40113: gap of 100 bp
40114 53470: contig of 13357 bp in length
53471 53570: gap of 100 bp
53571 69956: contig of 16386 bp in length
69957 70056: gap of 100 bp
70057 85519: contig of 15463 bp in length
85520 85619: gap of 100 bp
85620 108587: contig of 22968 bp in length
108588 108687: gap of 100 bp
108688 134650: contig of 25963 bp in length
134651 134750: gap of 100 bp
134751 169979: contig of 35229 bp in length.
FEATURES
source
1. .169979
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-22C7"
/clone_1lb="RP11-22C7"
/clone_1lb="RP11-22C7" Human Male BAC"
1. .1339
/note="assembly_fragment"
1440. .2542
/note="assembly_fragment"
2643. .4104
/note="assembly_fragment"
4205. .5645
/note="assembly_fragment"
5746. .10797
/note="assembly_fragment"
10898. .18365
/note="assembly_fragment"
18466. .28722
/note="assembly_fragment"
28823. .40013
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
40114. .53470
/note="assembly_fragment"
53571. .69956
/note="assembly_fragment"
70057. .85519
/note="assembly_fragment"

misc_feature 85620..108587 /note="assembly-fragment"
misc_feature 108688..134650 /note="assembly-fragment"
misc_feature 134751..169979 /note="assembly-fragment"
BASE COUNT 53870 a 34436 c 33100 g 47270 t 1303 others
ORIGIN
Query Match 69.7%; Score 20.2; DB 2; Length 169979;
Best Local Similarity 88.0%; Pred. No. 92;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 tctgcagccatgctagccccgc 29
|||||
Db 73881 TCTGCAGCCATGCTGGCAGCCGC 73905
RESULT 5
AC010811 165412 bp DNA linear HTG 09-FEB-2002
LOCUS Homo sapiens chromosome 11 clone RP11-3E17 map 11, *** SEQUENCING
IN PROGRESS ***, 5 unordered pieces.
AC010811 AC010811.10 GI:18642864
HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 165412)
TITLE Homo sapiens chromosome 11, clone RP11-3E17
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 165412)
REFERENCE
AUTHORS Balreen,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukngalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 9, 2002 this sequence version replaced g1:18543112.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: l2720
Center clone name: 3_E_17
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

* be preserved.
* 1 12794: contig of 12794 bp in length
* 12795 12894: gap of 100 bp
* 12895 68988: contig of 56094 bp in length
* 68989 69088: gap of 100 bp
* 69089 133991: contig of 64903 bp in length
* 133992 134091: gap of 100 bp
* 134092 162167: contig of 28076 bp in length
* 162168 162267: gap of 100 bp
* 162268 165412: contig of 3145 bp in length.
FEATURES
source
1..165412
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-3E17"
/clone_1lb="RPC1-11 Human Male BAC"
BASE COUNT 48029 a 30561 c 34765 g 51633 t 424 others
ORIGIN
Query Match 69.0%; Score 20; DB 2; Length 165412;
Best Local Similarity 82.1%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 atatctgacccatgctagccccgc 29
|||||
Db 29415 AGATCTGCAGCCATGACTTGGCCCTGC 29442
RESULT 6
AC068549 184018 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 11 clone RP11-559I10, WORKING DRAFT
DEFINITION SEQUENCE, 21 unordered pieces.
AC068549 AC068549.2 GI:7940407
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 184018)
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 184018)
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On May 19, 2000 this sequence version replaced g1:7684582.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0559I10
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175110 bases at least Q40
Consensus quality: 177590 bases at least Q30
Insert size: 20400; agarose-fp
Insert size: 182018; sum-of-contigs
Quality coverage: 4.34 in Q20 bases; agarose-fp


```
Quality coverage: 4.93 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1630: contig of 1630 bp in length
* 1631 1730: gap of unknown length
* 1731 3332: contig of 1602 bp in length
* 3333 3432: gap of unknown length
* 3433 5871: contig of 2439 bp in length
* 5872 5971: gap of unknown length
* 5972 8776: contig of 2805 bp in length
* 8777 8876: gap of unknown length
* 8877 11578: contig of 2702 bp in length
* 11579 11678: gap of unknown length
* 11679 15035: contig of 3357 bp in length
* 15036 15135: gap of unknown length
* 15136 17606: contig of 2471 bp in length
* 17607 17706: gap of unknown length
* 17707 22355: contig of 4649 bp in length
* 22356 22455: gap of unknown length
* 22456 26204: contig of 3749 bp in length
* 26205 26304: gap of unknown length
* 26305 31274: contig of 4970 bp in length
* 31275 31374: gap of unknown length
* 31375 39128: contig of 7754 bp in length
* 39129 39228: gap of unknown length
* 39229 46971: contig of 7743 bp in length
* 46972 47071: gap of unknown length
* 47072 57776: contig of 10705 bp in length
* 57777 57876: gap of unknown length
* 57877 66843: contig of 8967 bp in length
* 66844 66943: gap of unknown length
* 66944 78240: contig of 11297 bp in length
* 78241 78340: gap of unknown length
* 78341 93291: contig of 14951 bp in length
* 93292 93391: gap of unknown length
* 93392 110045: contig of 16653 bp in length
* 110045 110145: gap of unknown length
* 110145 124100: contig of 13956 bp in length
* 124101 124200: gap of unknown length
* 124201 140947: contig of 16747 bp in length
* 140948 141047: gap of unknown length
* 141048 158424: contig of 17377 bp in length
* 158425 158525: gap of unknown length
* 158525 184018: contig of 25494 bp in length.
*
FEATURES
source
1. 184018
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="11"
  /clone="RP11-559I10"
  1. 1630
misc_feature
  /note="assembly_name:Contig9"
  1731. 3332
misc_feature
  /note="assembly_name:Contig10"
  clone_end:SP6
  vector_side:left"
  3433. 5871
misc_feature
  /note="assembly_name:Contig11"
  5972. 8776
misc_feature
  /note="assembly_name:Contig12"
  8877. 11578
misc_feature
  /note="assembly_name:Contig13"
  11679. 15035
misc_feature
  /note="assembly_name:Contig14"
  15136. 17606
  /note="assembly_name:Contig15"
```

TITLE
 JOURNAL
 COMMENT
 Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Submitted (23-DEC-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 13, 2001 this sequence version replaced gl:15290826.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L11842
 Center clone name: 559_I_10

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 1 47886: contig of 47886 bp in length
 * 47887 47986: gap of 100 bp
 * 47987 57858: contig of 9872 bp in length
 * 57859 57958: gap of 100 bp
 * 57959 181028: contig of 123070 bp in length
 * 181029 181128: gap of 100 bp
 * 181129 185070: contig of 3942 bp in length.
 FEATURES
 source
 1.185070
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-559I10"
 /clone_lib="RPCI-11 Human Male BAC"
 BASE COUNT 61170 a 39478 c 35469 g 48576 t 377 others
 ORIGIN
 Query Match 69.0%; Score 20; DB 2; Length 185070;
 Best Local Similarity 82.1%; Pred. No. 1.1e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 2 atatctgcagcagctagcccccgc 29
 Db 41304 AGATCTGCAGCGCATGCTGGCCCTGC 41277
 RESULT 8
 AC073604_0/c
 WPCOMMENT
 Sequence split into 4 fragments LOCUS AC073604 Accession AC073604
 Fragment Name Begin End
 AC073604_0 1 110000
 AC073604_1 100001 210000
 AC073604_2 200001 310000
 AC073604_3 300001 352375
 LOCUS AC073604 352375 bp DNA linear HTG 13-OCT-2001
 DEFINITION Mus musculus clone RP23-93N5, WORKING DRAFT SEQUENCE, 37 unordered
 pieces.
 ACCESSION AC073604
 VERSION AC073604.3 GI:16117560
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE
 ORGANISM
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 1 (bases 1 to 352375)
 McCombie, W.R., Baker, J.P., Balija, V., Dedhia, N.N., de la
 Bastide, M., Kuit, K., King, L., Kirchoff, K.A., Miller, B.,
 Nascimento, L.V., O'Shaughnessy, A.L., Preston, R.R., Santos, L.,
 Shah, R.S., Spiegel, L.A., Palmer, L., Yang, C. and Zutavern, T.
 Mouse Genomic Sequence
 JOURNAL
 Unpublished
 TITLE
 2 (bases 1 to 352375)
 McCombie, W.R.
 REFERENCE
 Direct Submission
 Submitted (26-JUN-2000) Lita Annenberg Hazen Genome Sequencing
 Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
 Harbor, NY 11724, USA
 On Oct 13, 2001 this sequence version replaced gl:9972342.
 ----- Genome Center
 Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
 Laboratory
 Center code: CSHL
 Web site: http://www.cshl.org/genseq
 Contact: mcombie@cshl.org
 ----- Project Information
 Center project name: RP23-93N5
 Center clone name: RP23-93N5

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 37 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 1 26474: contig of 26474 bp in length
 * 26475 26658: gap of unknown length
 * 26659 49102: contig of 22444 bp in length
 * 49103 49286: gap of unknown length
 * 49287 69803: contig of 20517 bp in length
 * 69804 69987: gap of unknown length
 * 69988 86432: contig of 16445 bp in length
 * 86433 86616: gap of unknown length
 * 86617 102623: contig of 16007 bp in length
 * 102624 102807: gap of unknown length
 * 102808 118074: contig of 15267 bp in length
 * 118075 118258: gap of unknown length
 * 118259 132661: contig of 14403 bp in length
 * 132662 132845: gap of unknown length
 * 132846 147006: contig of 14161 bp in length
 * 147007 147190: gap of unknown length
 * 147191 159338: contig of 12148 bp in length
 * 159339 159522: gap of unknown length
 * 159523 171269: contig of 11747 bp in length
 * 171270 171452: gap of unknown length
 * 171453 182751: contig of 11299 bp in length
 * 182752 182934: gap of unknown length
 * 182935 194031: contig of 11097 bp in length
 * 194032 194214: gap of unknown length
 * 194215 205309: contig of 11095 bp in length
 * 205310 205492: gap of unknown length
 * 205493 216395: contig of 10903 bp in length
 * 216396 216578: gap of unknown length
 * 216579 227232: contig of 10654 bp in length
 * 227233 227416: gap of unknown length
 * 227417 236779: contig of 9364 bp in length
 * 236780 236962: gap of unknown length
 * 236963 245303: contig of 8341 bp in length
 * 245304 245486: gap of unknown length
 * 245487 253010: contig of 7524 bp in length
 * 253011 253193: gap of unknown length
 * 253194 260688: contig of 7495 bp in length

* 260689 260871: gap of unknown length
* 260872 260884: contig of 7213 bp in length
* 268085 268267: gap of unknown length
* 268268 275138: contig of 6871 bp in length
* 275139 275321: gap of unknown length
* 275322 281763: contig of 6442 bp in length
* 281764 281946: gap of unknown length
* 281947 288236: contig of 6290 bp in length
* 288237 288419: gap of unknown length
* 288420 294677: contig of 6258 bp in length
* 294678 294861: gap of unknown length
* 294861 301014: contig of 6154 bp in length
* 301015 301197: gap of unknown length
* 301198 306779: contig of 5582 bp in length
* 306780 306962: gap of unknown length
* 306963 312501: contig of 5539 bp in length
* 312502 312684: gap of unknown length
* 312685 317929: contig of 5245 bp in length
* 317930 318112: gap of unknown length
* 318113 323028: contig of 4916 bp in length
* 323029 323211: gap of unknown length
* 323212 328092: contig of 4881 bp in length
* 328093 328275: gap of unknown length
* 328276 332714: contig of 4439 bp in length
* 332715 332897: gap of unknown length
* 332898 336810: contig of 3913 bp in length
* 336811 336993: gap of unknown length
* 336994 340782: contig of 3789 bp in length
* 340783 340965: gap of unknown length
* 340966 344707: contig of 3742 bp in length
* 344708 344890: gap of unknown length
* 344891 348401: contig of 3511 bp in length
* 348402 348584: gap of unknown length
* 348585 352048: contig of 3464 bp in length
* 352049 352231: gap of unknown length
* 352232 352375: contig of 144 bp in length.
Location/Qualifiers
source
1. 352375
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-93N5"
BASE COUNT 92469 a 84017 c 82996 g 86234 t 6659 others
ORIGIN
Query Match 68.3%; Score 19.8; DB 2; Length 110000;
Best Local Similarity 91.3%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 3 tatctgcagccatgctagcc 25
|||||
Db 60510 TATCTGCAGCCTGGCCTAGGCC 60488
RESULT 9
AC094266/c 115062 bp DNA linear HTG 20-DEC-2001
LOCUS Rattus norvegicus clone CH230-3L5, *** SEQUENCING IN PROGRESS ***
DEFINITION 47 unordered pieces.
AC094266
AC094266.2 GI:17940982
VERSION HTG; HTGS_PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 115062)
AUTHORS Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Homsif,F., Howard,S., Huber,J., Huiyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisege,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshlari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 115062)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi.15636609.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAHL
Center clone name: CH230-3L5
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 96225 bases at least Q40
Consensus quality: 105112 bases at least Q30
Consensus quality: 112446 bases at least Q20
Estimated insert size: 93083; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 47 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 12330: contig of 12330 bp in length
* 12331 12430: gap of unknown length

*	12431	18306:	contig of 5876 bp in length
*	18407	18406:	gap of unknown length
*	18307	22841:	contig of 4435 bp in length
*	22842	22941:	gap of unknown length
*	22942	26470:	contig of 3529 bp in length
*	26471	26570:	gap of unknown length
*	26571	32226:	contig of 5656 bp in length
*	32227	33236:	gap of unknown length
*	32327	35350:	contig of 3024 bp in length
*	35351	35450:	gap of unknown length
*	35451	38659:	contig of 3209 bp in length
*	38660	38759:	gap of unknown length
*	38760	42446:	contig of 3687 bp in length
*	42447	42546:	gap of unknown length
*	42547	45127:	contig of 2581 bp in length
*	45128	45227:	gap of unknown length
*	45228	46887:	contig of 1660 bp in length
*	46888	46987:	gap of unknown length
*	46988	50671:	contig of 3684 bp in length
*	50672	50771:	gap of unknown length
*	50772	52400:	contig of 1629 bp in length
*	52401	52500:	gap of unknown length
*	52501	55326:	contig of 2826 bp in length
*	55327	55426:	gap of unknown length
*	55427	58414:	contig of 2988 bp in length
*	58415	58514:	gap of unknown length
*	58515	61460:	contig of 2946 bp in length
*	61461	61560:	gap of unknown length
*	61561	63526:	contig of 1966 bp in length
*	63527	63626:	gap of unknown length
*	63627	65784:	contig of 2158 bp in length
*	65785	65884:	gap of unknown length
*	65885	68062:	contig of 2178 bp in length
*	68063	68162:	gap of unknown length
*	68163	69962:	contig of 1800 bp in length
*	69963	70062:	gap of unknown length
*	70063	72736:	contig of 2674 bp in length
*	72737	72836:	gap of unknown length
*	72837	74716:	contig of 1880 bp in length
*	74717	74816:	gap of unknown length
*	74817	76796:	contig of 1980 bp in length
*	76797	76896:	gap of unknown length
*	76897	78076:	contig of 1180 bp in length
*	78077	78176:	gap of unknown length
*	78177	80644:	contig of 2468 bp in length
*	80645	80744:	gap of unknown length
*	80745	82356:	contig of 1612 bp in length
*	82357	82456:	gap of unknown length
*	82457	84035:	contig of 1579 bp in length
*	84036	84135:	gap of unknown length
*	84136	86753:	contig of 2618 bp in length
*	86754	86853:	gap of unknown length
*	86854	88307:	contig of 1454 bp in length
*	88308	88407:	gap of unknown length
*	88408	89643:	contig of 1236 bp in length
*	89644	89743:	gap of unknown length
*	89744	91020:	contig of 1277 bp in length
*	91021	91120:	gap of unknown length
*	91121	92656:	contig of 1576 bp in length
*	92697	92796:	gap of unknown length
*	92797	94248:	contig of 1452 bp in length
*	94249	94348:	gap of unknown length
*	94349	95368:	contig of 1020 bp in length
*	95369	95468:	gap of unknown length
*	95469	96603:	contig of 1135 bp in length
*	96604	96703:	gap of unknown length
*	96704	98728:	contig of 2025 bp in length
*	98729	98828:	gap of unknown length
*	98829	99950:	contig of 1122 bp in length
*	99951	100050:	gap of unknown length
*	100051	101229:	contig of 1179 bp in length
*	101230	101329:	gap of unknown length
*	101330	102503:	contig of 1174 bp in length

	FEATURES	source
*	102504	102603: gap of unknown length
*	102604	103796: contig of 1193 bp in length
*	103797	103896: gap of unknown length
*	103897	105091: contig of 1195 bp in length
*	105092	105191: gap of unknown length
*	105192	106200: contig of 1009 bp in length
*	106201	106300: gap of unknown length
*	106301	107720: contig of 1420 bp in length
*	107721	107820: gap of unknown length
*	107821	109208: contig of 1388 bp in length
*	109209	109308: gap of unknown length
*	109309	110945: contig of 1637 bp in length
*	110946	111045: gap of unknown length
*	111046	112361: contig of 1316 bp in length
*	112362	112461: gap of unknown length
*	112462	113735: contig of 1274 bp in length
*	113736	113835: gap of unknown length
*	113836	115062: contig of 1227 bp in length.
	Location/Qualifiers	
	1..115062	
	/organism="Rattus norvegicus"	
	/db_xref="taxon:10116"	
	/clone="CH230-3L5"	
BASE COUNT	33325 a	20155 c 21211 g 35735 t 4636 others
ORIGIN		

Query Match	68.38;	Score 19.8;	DB 2;	Length 115062;
Best Local Similarity	91.38;	Pred. No. 1.4e+02;		
Matches 21; Conservative	0;	Mismatches 2;	Indels 0;	Caps 0;

QY	4	atctgcagccatgctagcccc	26
Db	76679	AACTGCAGCCACGCGTAGGCCCC	76657

RESULT	10			
AC079643/c				
LOCUS				
DEFINITION	AC079643	151733 bp	DNA	linear
	Mus musculus	clone RP23-79H19,	WORKING DRAFT	SEQUENCE, 17 unordered
	pieces.			
ACCESSION	AC079643			
VERSION	AC079643.3	GI:14595775		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			

REFERENCE
AUTHORS
McCombie, W.R., Baker, J.P., Bahret, A., Yang, C., Ballja, V.,
1 (bases 1 to 151733)
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Dedhia, N.N., de la Bastide, M., Kuit, K., King, L., Kirchoff, K.A., Miller, B., Nascimento, L.U., O'Shaughnessy, A.L., Preston, R.R., Rodriguez, S., Santos, L., Shah, R.S., Spiegel, L.A., Palmer, L., Vll, M.D. and Zutavern, T.

TITLE	Mouse Genomic Sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 151733)

TITLE Direct Submission
JOURNAL Submitted (06-SEP-2000) Lita Annenberg Hazen Genome Sequencing

COMMENT On Jul 4, 2001 this sequence version replaced g1:10280742.

----- **Genome Center**

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor Laboratory

Center code: CSHL

web site: <http://www.cshl.org/genseq>

Contact: mccomb1@cshl.org

----- Project Information

Center project name: RP23-79H19

center clone name: RP23-79H19

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 24049: contig of 24049 bp in length
* 24050 24488: gap of unknown length
* 24489 41383: contig of 16895 bp in length
* 41384 41822: gap of unknown length
* 41823 54780: contig of 12958 bp in length
* 54781 55219: gap of unknown length
* 55220 67210: contig of 11991 bp in length
* 67211 67649: gap of unknown length
* 67650 77949: contig of 10300 bp in length
* 77950 78387: gap of unknown length
* 78388 88617: contig of 10230 bp in length
* 88618 89055: gap of unknown length
* 89056 96693: contig of 7638 bp in length
* 96694 97131: gap of unknown length
* 97132 103567: contig of 6436 bp in length
* 103568 104005: gap of unknown length
* 104006 110334: contig of 6329 bp in length
* 110335 110772: gap of unknown length
* 110773 116832: contig of 6060 bp in length
* 116833 117270: gap of unknown length
* 117271 123015: contig of 5745 bp in length
* 123016 123453: gap of unknown length
* 123454 129043: contig of 5590 bp in length
* 129044 129481: gap of unknown length
* 129482 134052: contig of 4571 bp in length
* 134053 134490: gap of unknown length
* 134491 138898: contig of 4408 bp in length
* 138899 139336: gap of unknown length
* 139337 143458: contig of 4122 bp in length
* 143459 143896: gap of unknown length
* 143897 147887: contig of 3991 bp in length
* 147888 148325: gap of unknown length
* 148326 151733: contig of 3408 bp in length.

FEATURES
Source location/Qualifiers

1.151733
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-79H19"
/clone_1lb="RPCI-23"

BASE COUNT 33878 a 37878 c 38717 g 34033 t 7227 others
ORIGIN

Query Match 68.3%; Score 19.8; DB 2; Length 151733;
Best Local Similarity 91.3%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 tatctgcagccatgctagcccc 25
|||||
Db 94130 TATCTGCAGCCTGGGCTAGGCC 94108

RESULT 11
AF321096/c 1560 bp DNA linear PLN 15-MAR-2001
LOCUS
DEFINITION Pichia pastoris PR-aminimidazolesuccinocarboxamide synthase (ADE1)
gene, complete cds.

ACCESSION AF321096
AF321096.1 GI:12744763

VERSION
KEYWORDS
SOURCE
ORGANISM

Pichia pastoris.
Pichia pastoris
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE 1 (bases 1 to 1560)

AUTHORS

Lin Cereghino,G.P., Lin Cereghino,J., Sunga,A.J., Johnson,M.A.,

Lim,M., Gleeson,M.A.G. and Gregg,J.M.

New selectable marker/auxotrophic host strain combinations for
molecular genetic manipulation of Pichia pastoris

Gene 263 (1-2), 159-169 (2001)

11223254

REFERENCE

2 (bases 1 to 1560)

Lin Cereghino,G.P., Lin Cereghino,J., Sunga,A.J., Johnson,M.A.,

Lim,M., Gleeson,M. and Gregg,J.M.

Direct Submission

Submitted (14-NOV-2000) Biological Sciences, University of the

Pacific, Stockton, CA, USA

FEATURES

Location/Qualifiers

1.1560

/organism="Pichia pastoris"

/db_xref="taxon:4922"

<371. >1285

/gene="ADE1"

/product="PR-aminimidazolesuccinocarboxamide synthase"

<371. >1285

/gene="ADE1"

371.1285

/gene="ADE1"

/function="biosynthetic enzyme necessary for production of

adenine"

/note="adel mutant has pink phenotype"

/codon_start=1

/product="PR-aminimidazolesuccinocarboxamide synthase"

/protein_id="AAK06766.1"

/db_xref="GI:12744764"

/translation="MSIVNTDIDGILPLIAKKVRIYAVDENLLFVATDRISADV

IMTNGIPDKGKILTLQSVFWDPLAPYIKNLVASNDKEFALLPSKLSSEKYSOLE

GRSLIVKKHRLIPLEAIVRGYITGSAMKEYKNSKTVHGAVENENLQESDAFPPIFT

PSTKAEQGEHDENISIEQAETVGKDICEKAVKAVELYSAAKNFALLKGIITADTKF

EFGLDENNELVLVDEVLTTPDSRFWNOKTYOVGKSQSESYDKOFLRDWLTANLNGREG

VAMDAEIAIKSKEYIEAYEATGKWA"

BASE COUNT

494 a 285 c 329 g 452 t

ORIGIN

Query Match 67.6%; Score 19.6; DB 8; Length 1560;
Best Local Similarity 84.6%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 atatctgcagccatgctagcccc 27
|||||

Db 1218 ATTTCTGCATGCATGCCTACGCCCTC 1193

RESULT 12
AX175474/c 11358 bp DNA linear PAT 03-JUL-2001
LOCUS
DEFINITION Sequence 3 from Patent WO0144443.
AX175474
VERSION AX175474.1 GI:14598816
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 11358)
AUTHORS White,J.A., Petkovich,P.M., Jones,G. and Ramshaw,H.
TITLE Cytochrome p450rai-2 and related proteins
JOURNAL Patent: WO 0144443-A 3 21-JUN-2001;
Cytochrome Inc. (CA)

FEATURES
source Location/Qualifiers

1.11358

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 2654 a 2894 c 3040 g 2770 t

ORIGIN

Query Match 67.6%; Score 19.6; DB 6; Length 11358;
Best Local Similarity 84.6%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 atctgcagccatgctagtagcccccgc 29
|||||
Db 760 ATCTGCAGCCAGGCGCCAGCCACG 735

RESULT 13
AE007115 15873 bp DNA linear BCT 27-APR-2001
LOCUS Mycobacterium tuberculosis CDC1551, section 201 of 280 of the
DEFINITION complete genome.
ACCESSION. AE007115 AE000516
VERSION AE007115.1 GI:13882670
KEYWORDS
SOURCE Mycobacterium tuberculosis CDC1551.
ORGANISM Mycobacterium tuberculosis CDC1551.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., Deboy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
TITLE Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 15873)
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., Deboy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
Gill,J., Mikula,A. and Bishai,W.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1. .15873
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
128. .877
/gene="MT2897"
128. .877
/gene="MT2897"
/note="similar to GB:U00096 SP:P76082 PID:1742269
PID:1742275 PID:1787659; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK47223.1"
/db_xref="GI:13882671"
/translation="MTDDLIDTDERVRTLTNRPQSRNALSAALDRFFAALADAE
ADDIDIVILTGADPVFCAGDLKELAGQTALPDISRPWPMTKPVIGAINGAAVTGG
LELAICYDILIASSEHARFADHARVGLLPTWGLSVRLPKQVGIGLARMSLTGDYLSA
TDALRAGLVEVVAHDQLPTARVVAASIVGNQNAVRLALASYHRIDESQTAAGLWL
EACAKOFRTSQDTIAANREAVLQGRGAQVR"
complement(956. .2038)
/gene="MT2898"
complement(956. .2038)
/gene="MT2898"
/note="similar to GI:1870167; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="sugar ABC transporter, ATP-binding protein"
/protein_id="AAK47224.1"

gene
CDS
/db_xref="GI:13882672"
/translation="MANQYSAVTQRYPGADAPTVDNLDLDADGEFLVLVGPSCGK
STTLRVLAGLEPIESGRISIGDVDVTHLPPRARDVAMVFQNYALYPMNTVAANMGFAL
RNAGMSRADTTRRRVLEVADMLFETDLDRKPAKLISGQRRQRYAVAGRAIVRRPVFCMD
EPLSNLDAKLRVSTRSQISGLQRLGTTTVYVTHDQVEAMTMGRVAVLKDGLQQVD
TPRALYDDPVNTFVATEFICAPAMNLDAAVAHGVRAPDLAIPVDPAAERVLGVVR
ESWDVASIGTPGSLTVHVELVEEFSEFVYATPVQGRWSSRAPRIVERTRDRTAVR
VGESLAIVPHSQEVRLENSRTETRLR"
complement(2031. .3083)
/gene="MT2899"
complement(2031. .3083)
/gene="MT2899"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="sugar ABC transporter, sugar-binding protein"
/protein_id="AAK47225.1"
/db_xref="GI:13882673"
/translation="MRRSSEPTCPSPCSTTDGGSISPSAVESLPLTCSAKLGWQK
ITSIRCWPTMSSTAITLCRMLARRCSTTRRGNRPAYPTADNPGQSSSTGVSRY
SAMSAPVDRTAGNADLISWTFQGPMAFGAYSQDKWTLTLEPATIAAGNFYRNSIH
GKGAAVANDIANEFATGILASAVASTGLAGITASARFDFGAPLPTGPDAAPACP
GGAGLAIPAKLSEERKVNALKFAEVTNPNTAFYSQOTGYLPVRKSAVDASERHYL
ADNPRARVALDQLPHTRTQDVARVFLPGDRIISAGLESIGRGADVTKTFTNIQRL
QVILDRQIMRKLHG"
complement(3342. .4169)
/gene="MT2900"
complement(3342. .4169)
/gene="MT2900"
/note="similar to SP:P10906 GB:U00039 PID:43248 PID:466587
GB:U00096; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="sugar ABC transporter, permease protein"
/protein_id="AAK47226.1"
/db_xref="GI:13882674"
/translation="MTPDRLRSSVGYAAMLLVVTLIAGPLFEVFTSEKQDPDIYAQ
TSWPLRWYPQNYRTATEQIPEWTFNRNSLITSVLAVVKEFTGLVLSAFGLVEFRPG
RTAVFLVITIAALMVPNOITVISNALISHLGRNTFAGIILPLAGVAFGFTPLMRNHL
SLPAELIEAARMMDGARWQOLLRLRVLPMSRPTMVAVGVIVVNEVNEYLMPFLMSDE
SVAPLPGLTFLQAEQVTNWCPVNAVTLMLPILLVFIALQROMIKGLTGAVKG"
complement(4166. .5077)
/gene="MT2901"
complement(4166. .5077)
/gene="MT2901"
/note="similar to SP:P10905 GB:U00039 PID:43247 PID:466588
GB:U00096; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="sugar ABC transporter, permease protein"
/protein_id="AAK47227.1"
/db_xref="GI:13882675"
/translation="MAAPQARARLSSKERVDYALFVYVGNVALLLFFVYRPLADN
IRLSFTDMNVSDPSARFVGLSNYTEWFTRSQDTRQIVFNATVETGAAVGSMVLGLALA
MLDRPLRGRNLVRSSTVFAPFVISAAGVLAQVFDPHFGLIQDLLRRIGVGVPDFY
QDARNALEWVTITYVKNLGYTFVYIYLAALQGVARDLLEAETDGASRWAVERRVLLP
QLRPTTFELSLITVLINSLOQVEDVINVMTRGPEGTGTTMYQYVYVETFRNFRAGYGA
TVATIMFLVLLAVTYQVRVMDRGQRQ"
complement(5164. .6474)
/gene="MT2902"
complement(5164. .6474)
/gene="MT2902"
/note="similar to SP:P28303 PID:146611 PID:396379
GB:U00096 PID:1790477; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="DNA-damage-inducible protein F, putative"
/protein_id="AAK47228.1"
/db_xref="GI:13882676"
/translation="MGRAGRGRTAQLALPALGVLAEPYLLFDIAVVGRLGAISLA
GLAIGSLVLGVSOATFLSYGTARAAARYGAGNRVAAYEGVQATWALGLGALVY
VVEATATPLVSAIASGDGITAALPMLRIAIIIGTPAIYLSLAGNGWLKRGVQDTVRPL

RVVAGFGSSALLCPILVYGMGLPRWGLTGSAVANLVCWLALLFAGALLAERVSL
RPDRAVLGAQLMMARDLIVRTLAFOVCYVSAAVAARFGAALAHOVLTQWGLLAL
VLDSLAIAGSLVGAALGADGAGAKAVAMRVTAFLSLAAGILAAHVLGSSVLPGLF
TDDRSVLAIGVPMWFMVYQLPEAGIVAVDGVLLGAGDAFMRTATVASALVGFLEPL
VWLSLAYGWLAGIWSGLGTFTIVRLRLEFVGWRAYSGRWAVTGAA"
gene
complement(6490. .7440)
/gene="MT2903"
CDS
complement(6490. .7440)
/gene="MT2903"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="DHH family protein"
/protein_id="AAK47229.1"
/db_xref="GI:13882677"
/translation="MDAVGAALLSAAARVGVCHVPDADTIGAGLALVLDGCGK
RVEVSFAAPATLPESLSLPGCHLLVRPEVMRDVLDVTVDIPSVDRLGALGDLDS
GRELLVIDHNASNDLFGTFANFIDPSADSTTTVAEILDAMGKPIDPRVAHCIVAGLAT
DTGSFRWASVRCYRLAARLVEIGVDNATVSRITLMDSHPTWLPILSRVLGSAQLVSEA
VGGRLVYVVVDNREWVAARSEEVESIVDIVRTQAEVAAVFEVPEHRWSVSMRAK
TVNLAAVASGFGGGHRLAAGYTTTGSIDDAVASLRALALG"
gene
complement(7475. .8026)
/gene="MT2904"
CDS
complement(7475. .8026)
/gene="MT2904"
/note="similar to SP:P32731 GB:Z18631 PID:580902
GB:AL009126; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosome-binding factor A"
/protein_id="AAK47230.1"
/db_xref="GI:13882678"
/translation="MADARARLAKRIAIVASAIYEIKDPLAGVTITDAKYTAD
LHDATVYTYVMGRTLHDEPNCAAAALERAKGLRTKVGAGTGVRTPTLTFTLDTI
SDSVHRMDELLARARADADLARVRVAKPAGEADPYRDNGSVAQSPAPGGLGIRTS
GPEAVEAPLTCCGDVDDDRPKE"
gene
complement(8026. .10728)
/gene="MT2905"
CDS
complement(8026. .10728)
/gene="MT2905"
/note="similar to SP:P04766 GB:X04399 PID:39954;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="translation initiation factor IF-2"
/protein_id="AAK47231.1"
/db_xref="GI:13882679"
/translation="MAAGKARVHELAKELGVTSEKYLARLSEQEYFKSASSTVEAPV
ARLRESFGSKPAPAKGTAKGPDKSLDKALDAIDMAAGNGKATAAPAKADS
GGAIVSPTTPAAPPEPTAVPPSPQAPHPGMAPGARPVPKPGIRTPRVGNPSSA
QPADRPIDRPAPRPGTARPGVPRPGASPGSMPPRGAVGARGPRPGARPGRC
APGAGSDAGGNGYRGVGAAPGTGFRGGGGGGGPGGAGAGAFGRPGAPRR
GRKSKROKROYDSMQAPVYGVRLPHNGETIRLARGASLSDFADKIDANPALVQA
LENLGEVMTATQSVGDETELLGSEMNYNQVVSPEDERELLESFDSLSTGEDEGEE
DLQVRPVVTVMGHVDHCKTRLDDTIRKANVREAEAGITTOHIGAYQVAVDLDSQRL
ITFIDTPGHEAFTAMRARGAKATDIALVVAADDGVMQTEAINHAAQADVPPIVAV
NKIDKEGADPAKIRGLTEYGLVPEEFGDTMFVDISAKOGTNIEALEAVLITADAA
LDLRANPDMEAOGVAIEAHIDRGRPVATVLRGTLRGDSVAGDAYGRVRRMVD
HGEDVEVALPSRPVOVIGFTSVPGAGDNFLVDEDRIRAOIADRRSARKRNALAA
KRISLEDLDSALKETSQLNLILKGNAGTVEALEALMGIQVDEVVLRVIDRGVGI
TETNVNLASASDAVITGFNVRAEGKATELASREGVEIRYSVIYQAIDEIEQALRGL
KPIYEENOLGRAEIRALFRSSKVLGATCLVTSGVMRNAKARLLRDNINVAENLSLA
SLRREKDDVTEVRDGFEGCGTLTGADIKEGDIESTELVQKERA"
gene
complement(10814. .11113)
/gene="MT2906"

Query Match 67.6%; Score 19.6; DB 1; Length 15873;
Best Local Similarity 84.6%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 atctgcagcatgctagggcccccgc 29
|||||
Db 6080 ATCCGACGACGCAAGCGCCGCCG 6105

RESULT 14
MSGY414A/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

MSGY414A
Mycobacterium tuberculosis sequence from clone y414a.
AD000007
AD000007.1 GI:1702972
Mycobacterium tuberculosis (clone: y414a) ds-DNA.
Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 40121)
Du, L.
Direct Submission
Submitted (11-OCT-1996) L. Du, Genome Therapeutics Corporation, 100
Beaver Street, Waltham, MA, USA, 02154 du@cric.com
GSDS:S:1004719
Location/Qualifiers
1. 40121
/organism="Mycobacterium tuberculosis"
/db_xref="taxon:1773"
/clone="y414a"

BASE COUNT 7156 a 12828 c 13251 g 6885 t 1 others
ORIGIN

Query Match 67.6%; Score 19.6; DB 1; Length 40121;
Best Local Similarity 84.6%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 atctgcagcatgctagggcccccgc 29
|||||
Db 5823 ATCCGACGACGCAAGCGCCGCCG 5798

RESULT 15
MTCY16B7
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

MTCY16B7
Mycobacterium tuberculosis H37Rv complete genome; segment 123/162.
Z81331 AL123456
Z81331.1 GI:3261650
Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 43430)
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E.,
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,
Squares, S., Squares, R., Sultson, J.E., Taylor, K., Whitehead, S. and
Barrell, B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature. 393 (6685), 537-544 (1998)
98295987
2 (bases 1 to 43430)
Parkhill, J.
Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:1648878.

Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes implemented in Tbpase (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

FEATURES
source
1. .43430
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
<1. .507
/note="fragment designated v002. Does not represent a physical clone"
58. .113
/note="hairpin loop with 38 bp stem"
complement(133. .696)
/gene="lppv"
complement(133. .696)
/gene="lppv"
/note="Rv2796c, (MTV002.61c-MTCY16B7.47), len 187; Probable lipoprotein similar to MTCY159.13c (220 aa) fasta score:opt: 169 z-score: 206.2 E(): 0.00058; 22.4% identity in196 aa overlap"
/codon_start=1
/transl_table=11
/product="lppv"
/protein_id="CAB03673.1"
/db_xref="GI:3261651"
/db_xref="SPTREMBL:P71655"
/translation="MRWPTAWLALVCMATGCGPSGHGTRAGEGRLSPEKVAELN PLRAKPLEDAKDQYRAAVTQLANITALLVPGLTWRTDMDTWCGEYETRAKAA YFMIVESGPIPDCKWLQAVQIVKDGVEQATGFGVMKNKPADHDVYFAGHGVEFKCS TOKAAVLTAQSDCRISRTDTPKPSPTP"
382. .>43430
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="Y16B7"
complement(696. .2384)
/gene="Rv2797c"
complement(696. .2384)
/gene="Rv2797c"
/note="Rv2797c, (MTCY16B7.46), len: 562. Unknown, containsPS00120 lipases, serine active site, similarity to 283863|MTCY159_14 Mycobacterium tuberculosis cosmid (403 aa) opt:537 z-score: 530.8 E(): 4.8e-22; 40.8% identity in 292aa overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2797c"
/protein_id="CAB03649.1"
/db_xref="GI:1648924"
/db_xref="SPTREMBL:P71654"
/translation="MPLTVADIDRNNAAQAVREVFHAASARAENVTEASRQLALSTIFA NSGKTAEBAAAHNAGIRRDIGLADNEALAVARAADRAADGIVKQSELALRHAAA AELTIDALINRVVPIGLRSTEAQWARTLAKOTELQAEIDAIMAEANVDEELASAVN MADGAPIPADSGRPVGEGLTPQLASDANEERLREERARLQAHLERLQAEYDQLSV RAARDYHNGILDGDVAGRLAALDELSAARGRLGELDAVDALSRAPETYLTOLOIPE DPNQOVLAAVAVGNPDITAAVNSVTVPGVSTRGALPGMVTEARDLRSEVIRQLNAAG KPAVAVTIAMGCVHPPNPRLTGSAGDLQMTMDGAHAGADLSRYLQOVRRANPSCG HLTVLGHSYGSLTASLALQDLDAQSAHPVNDVVFYGSPLGLTYSQAQLGLDGHAYVM QAPHDLITNLVAPLALPHGWLDPYLTLPGETELSSQAGFDGSIWRDGVYAHGDYPRS

misc_feature
complement(1176. .1205)
/gene="Rv2797c"
/note="PS00120 lipases, serine active site"
complement(2388. .2714)
/gene="Rv2798c"
complement(2388. .2714)
/gene="Rv2798c"
/note="Rv2798c, (MTCY16B7.45), len: 108. Some similarity to Rv0965c(MTCY10D7.09) (139 aa, E(): 1.6e-06, 38.9% identity in 90 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2798c"
/protein_id="CAB03648.1"
/db_xref="GI:1648923"
/db_xref="SPTREMBL:P71653"
/translation="MFOISPEQWMHSAQVTTQEGGLAVGHLSSDYMQAQFGWOGA SAMALNAKMDDWLDASRALLTRIGDHFLQLEAIAIQHAAEAERQAALAQVGSADVV AGPRGV"
2845. .3474
/gene="Rv2799"
2845. .3474
/gene="Rv2799"
/note="Rv2799, (MTCY16B7.44c), len: 209. Unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2799"
/protein_id="CAB03647.1"
/db_xref="GI:1648922"
/db_xref="SPTREMBL:P71652"
/translation="MYTPGKPPRAGGVFTRVRLIGGLALTAVVVGVGWGIGP PAPTGDAVQLRSTAAPMSTMTSPVATTDPSFDPDPCRDIPFDVIRGLAYTPPEA EEGLRCHFDAGNYOMAVEPIIWRITAOQLPPDAIETTIAGHRAAQYVWRKPTYHNSFW YSSCMVTEKTSYGVIIQOSLFYSTVSEPDVDCPSTNLQRANDLVYYRF"
3493. .5142
/gene="Rv2800"
3493. .5142
/gene="Rv2800"
/note="Rv2800, (MTCY16B7.43c), len: 549. Some similarity to G216374 GLUTARYL 7-ACA ACYLASE precursor (634 aa), fasta scores, opt: 362, E(): 1e-14, (26.7% identity in 589 aa overlap) and to D90903|D90903_22 Synechocystis sp; PCC6803 complete (535 aa) opt: 455 z-score: 494.2 E(): 5.3e-20; 29.8% identity in 503 aa overlap. Some similarity to Rv1835c, (MTCY1A11.08)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2800"
/protein_id="CAB03646.1"
/db_xref="GI:1648921"
/db_xref="SPTREMBL:P71651"
/translation="MSTTSARPERPKRALTGRVGGQALGSLGPRATTRYTVGHVR VPMRDGVQLVADHYAPATSQPVGTLVGRPYGRREPFSLVFARIYAARGYHVLQSVR GTFSGGVEEPMVNEAADGADTVAMRQPMETGREGTIGLPIYLGFTQWALLHDPPE LAAAVITVGPHPDEFASVWGTSFTVNDFLGMSDLVSHQEDPGRIRAGIRQLTAPRRA RTAATPLDGESARTLLGTGAPWFESVNEHTDRDDPEWDRLRFPALDRVQVPLLVGSG WODIFLRQTLQOYRHLRDRGVHVALTVGPWTHOTMLTKGLATGARESLDWLDAHLGRA PALRPSRVRFVVTGGWRHLPMWPATTERAWYLQPGGRLGESAPASGTPPATFRYHP ADPTPTGGPRLSSNGYRDDSRLATRADVLCFTGAPLTHDCVHGNPVVELVHSSDN PYVDVEYRVSEVDAKGRSRNVSDBGYRRLGDAPELVRVELDATAHFRADSRIRVLIAG SWFPRYARNIGTPEPILGTGROLKPATHAVHFGSRLLLPVG"
complement(5244. .5600)
/gene="Rv2801c"
complement(5244. .5600)
/gene="Rv2801c"
/note="Rv2801c, (MTCY16B7.42), len: 118, similar to B. subtilis sp|P96622|P96622 YDCE PROTEIN (116 aa) fasta scores, opt: 194 z-score: 275.5 E(): 8e-08; 33.3% identity in 117 aa overlap; PEMK_ECOLI P13976 pemk protein (133 aa), fasta scores, opt: 153, E(): 0.00055, (35.6% identity in 118 aa overlap), also MTCY39.28 (36.8% identity in 117 aa overlap)"


```
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2801c"
/protein_id="CAB03645.1"
/db_xref="GI:1648920"
/db_xref="SPTREMBL:P71650"
/translation="MMRGEIWIQVDLDPARGSEANNQRPVAVVSNDRANATATRLGRG
VITVVPVTSNIAKYVPQVLLSATTTGLQVDCKAQAEQIRSIATERLLRPIGRVSAAE
LAQLDEALKLHLDLWS"
complement(5857..6900)
/gene="Rv2802c"
/complement(5857..6900)
/gene="Rv2802c"
/note="Rv2802c, (MTCY16B7.41), len: 347. Unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv2802c"
/protein_id="CAB03674.1"
/db_xref="GI:3261652"
/db_xref="SPTREMBL:P71649"
/translation="MARQPLEQRVARAAQAALARQRFVSAIDVLLGLGLAPSHVDQW
```

```
Query Match      67.6%; Score 19.6; DB 1; Length 43430;
Best Local Similarity 84.6%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY      4 atctgcagccatgctagcccccgc 29
      ||| ||||| ||| ||||| |||
Db 38302 ATCCGCAGCCACGGCAAGCGCCGCGC 38327
```

Search completed: August 31, 2002, 18:03:50
Job time: 15683 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 13:43:18 ; Search time 824.08 Seconds
(without alignments)
60.420 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatatctgcagccatggttagggccccgcg 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	29	100.0	29	AAS17545	Human acetylcholin
C 2	19.6	67.6	11358	AAH22434	P450RAI-2 containi
3	19.2	66.2	375	AAF66350	Novel human polynu
4	19.2	66.2	657	AAH52128	Human AFP protein
5	19.2	66.2	1921	AAS76387	DNA encoding novel
6	19.2	66.2	2390	AAS45000	cDNA encoding novel
C 7	19.2	66.2	3023	AAS77493	DNA encoding novel
C 8	19.2	66.2	3684	AAS80007	DNA encoding novel
9	19	65.5	934	AAL36719	Human musculoskele

C 10	19	65.5	934	22	AAL36720	Human musculoskele
11	19	65.5	1173	21	AAA37841	Mouse angiopoietin
12	19	65.5	1223	22	AAS22675	Human CDNA encodin
13	19	65.5	1260	22	AAS22439	Human CDNA encodin
C 14	18.6	64.1	702	22	AAH08668	Human CDNA clone (
C 15	18.6	64.1	703	22	AAH06512	Human CDNA clone (
C 16	18.6	64.1	1783	22	AAH18331	Human CDNA sequenc
C 17	18.6	64.1	2368	22	AAH16240	Human CDNA homolog
18	18.6	64.1	11849	20	AAS223905	Human LOBO homolog
C 19	18.4	63.4	6177	23	ABL11819	Drosophila melanog
C 20	18.4	63.4	18294	23	ABL11818	Drosophila melanog
C 21	18.2	62.8	51	22	AAL31467	Human SNP oligonuc
C 22	18.2	62.8	204	19	AAH10671	Human biallelic po
23	18.2	62.8	2109	22	AAS02049	DNA encoding molec
C 24	18.2	62.8	3108	14	AAQ46248	Aspergillus niger
C 25	18.2	62.8	3108	14	AAQ48459	catR gene. Asperg
C 26	18.2	62.8	12697	23	ABL10888	Drosophila melanog
C 27	18	62.1	440	22	AAK04999	Human brain expres
C 28	18	62.1	440	22	AAK17722	Human brain expres
C 29	18	62.1	440	22	AAK30534	Human bone marrow
C 30	18	62.1	440	22	AAK43544	Human bone marrow
C 31	18	62.1	440	22	AAI36476	Probe #5162 used t
C 32	18	62.1	440	22	AAI49598	Probe #18284 used
C 33	18	62.1	440	22	AAI04893	Probe #4884 used t
C 34	18	62.1	440	22	AAI09872	Probe #9863 used t
35	18	62.1	1560	22	AAF31049	Abscisic acid-acti
C 36	18	62.1	18286	22	AAK70889	Human immune/haema
C 37	18	62.1	18286	22	AAK84104	Human immune/haema
C 38	18	62.1	18287	22	AAK70892	Human immune/haema
C 39	18	62.1	18287	22	AAK84107	Human immune/haema
C 40	17.8	61.4	534	22	AAH67239	C glutamicum codin
41	17.8	61.4	595	22	AAK59851	Human immune/haema
42	17.8	61.4	633	21	AAF08384	Fusarium venenatum
C 43	17.8	61.4	1113	14	AAQ41298	Human BMP-3 gene.
C 44	17.8	61.4	1774	13	AAQ32852	BMP3. Rattus ratt
C 45	17.8	61.4	1794	10	AAH92276	Region of clone la

ALIGNMENTS

RESULT 1
AAS17545
ID AAS17545 standard; DNA; 29 BP.
XX
AC AAS17545;
XX
DT 25-FEB-2002 (first entry)
XX
DE Human acetylcholinesterase, PCR primer AChE-Nco.

XX
KW Human; acetylcholinesterase; AChE; PCR primer; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ss.

OS Homo sapiens.
OS Synthetic.

XX WO200171014-A2.

PD 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US08468.

XX 17-MAR-2000; 2000US-190440P.

PA (MORT/) MOR T.

PA (SORE/) SOREQ H.

PA (ARNT/) ARNTZEN C.

PA (MASO/) MASON H.

PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX

PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 4; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents PCR primer AChE-Nco used to amplify
CC human AChE as described in the method of the invention.
XX
SQ Sequence 29 BP; 5 A; 11 C; 8 G; 5 T; 0 other;

Query Match 100.0%; Score 29; DB 24; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.0066;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatattgcagccatggctagggcccccgc 29
Db ||||||||||||||||||||||||||||
1 gatattgcagccatggctagggcccccgc 29

RESULT 2
AAH22434/c
ID AAH22434 standard; DNA; 11358 BP.
XX
AC AAH22434;
XX
DT 22-AUG-2001 (first entry)
XX
DE P450RAI-2 containing human genomic DNA (clone NH0493L16).
KW Cytochrome P450; P450RAI-2; brain; retinoic acid; cancer; dysplasia;
KW autoimmune; dermatological; cytostatic; antiinflammatory; antiseborrheic;
KW antipsoriatic; immunosuppressive; ds.
XX
OS Homo sapiens.
XX
PN WO200144443-A2.
XX
PD 21-JUN-2001.
XX
PF 15-DEC-2000; 2000WO-CA01493.
XX
PR 16-DEC-1999; 99US-0171110.
PR 27-JAN-2000; 2000US-0178314.
XX
PA (CYTO-) CYTOCHROME INC.
XX

PI White JA, Petkovich PM, Jones G, Ramshaw H;
XX WPI; 2001-390242/41.
DR P-PSDB; AAB85151.
DR
XX Novel P450 protein useful for metabolizing retinoic acid for treating
PT cancer, dysplasia, an autoimmune or dermatological disease -
XX
PS Example 1; Page 155-157; 174pp; English.
XX
CC The present invention provides a novel all-trans-RA metabolising
CC cytochrome P450, P450RAI-2, that is predominantly expressed in the
CC brain. This novel cytochrome P450 is useful for metabolising retinoic
CC acid in a cell or organism, for screening drugs for their effect of
CC protein activity, oxidizing a retinoid, screening an agent for its effect
CC on protein activity. The P450RAI-2 polypeptide, nucleic acid and host
CC cells containing them are useful for treating cancer, dysplasia, an
CC autoimmune or dermatological disease. A drug which has an effect on the
CC expression of P450RAI-2 is used to inhibit retinoic acid metabolism in
CC the treatment of cancer, actinic keratosis, oral leukoplakia, a secondary
CC head and/or neck tumour, a non-small cell lung carcinoma, a basal cell
CC carcinoma, skin cancer, and a premalignancy associated actinic keratosis,
CC acne, psoriasis, ichthyosis, and/or preferably acute promyelocytic
CC leukemia. The present sequence represents a human genomic sequence from
CC GenBank (Accession#, AC007002 clone name NH0493L16). This sequence
CC comprises the DNA encoding the novel cytochrome P450, P450RAI-2.
XX
SQ Sequence 11358 BP; 2654 A; 2894 C; 3040 G; 2770 T; 0 other;

Query Match 67.6%; Score 19.6; DB 22; Length 11358;
Best Local Similarity 84.6%; Pred. No. 87;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 atctgcagccatggctagggcccccgc 29
Db ||||||||||||| ||| ||| ||| ||
760 ATCTGCAGCCAGGCCAGGCCAGGCCAGC 735

RESULT 3
AAF66350
ID AAF66350 standard; cDNA; 375 BP.
XX
AC AAF66350;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2106.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX

PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 850; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 375 BP; 71 A; 140 C; 105 G; 58 T; 1 other;

Query Match 66.2%; Score 19.2; DB 22; Length 375;
Best Local Similarity 87.5%; Pred. No. 94;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 atctgcagccatggctaggccccc 27
|||||
Db 241 atctgcagccatggcattggacccc 264
|||||

RESULT 4
AAH52128
ID AAH52128 standard; cDNA; 657 BP.
XX
AC AAH52128;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human AFP protein encoding cDNA sequence SEQ ID NO:71.
XX
KW Human; secreted protein; secretion; bacterial cell; fungal cell;
KW eukaryotic cell; fusion protein; maltose binding protein;
KW immunoglobulin constant region; polyhistidine tag; ss.
XX
OS Homo sapiens.
XX
PN WO200129221-A2.
XX
PD 26-APR-2001.
XX
PF 20-OCT-2000; 2000WO-US29052.
XX
PR 20-OCT-1999; 99US-0160712.
XX
PA (ZYMO) ZYMOGENETICS INC.
PI Conklin DC, Yee DP;
XX
DR WPI; 2001-300340/31.
DR P-PSDB; AAG81277.
XX
PT Isolated polypeptide for directing secretion of proteins of interest
PT from a host cell including, e.g. bacteria, includes contiguous amino
PT acid residues of polypeptide with specified amino acids -
XX
PS Claim 9; Page 174-176; 617pp; English.
XX
CC AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242
CC to AAG81453. The secreted proteins can be used for directing the

CC secretion of proteins of interest from a host cell including bacteria,
CC fungal cells, and cultured higher eukaryotic cells. The present invention
CC also describes fusion proteins, where a secreted protein of the invention
CC is operably linked via a peptide bond or peptide linker to a second
CC protein selected from the group consisting of maltose binding protein,
CC an immunoglobulin constant region, a polyhistidine tag and a peptide
CC given in AAG81453.
XX
SQ Sequence 657 BP; 132 A; 214 C; 207 G; 104 T; 0 other;

Query Match 66.2%; Score 19.2; DB 22; Length 657;
Best Local Similarity 87.5%; Pred.No. 99;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 atctgcagccatggctaggccccc 27
|||||
Db 571 atctgcagccatggcattggacccc 594
|||||

RESULT 5
AAS76387
ID AAS76387 standard; cDNA; 1921 BP.
XX
AC AAS76387;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #12191.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG12200.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 12191; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1921 BP; 445 A; 558 C; 540 G; 378 T; 0 other;

Query Match 66.2%; Score 19.2; DB 23; Length 1921;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 atctgcagccatggctaggccccc 27
|||||
Db 988 atctgcagccatggctaggccccc 1011

RESULT 6
AAS45000
ID AAS45000 standard; cDNA: 2390 BP.
XX
AC AAS45000;
XX
DT 18-DEC-2001 (first entry)
XX
DE cDNA encoding novel human secretory protein, Seq ID No 81.
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200166689-A2.
XX
PD 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US04942.
XX
PF 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR WPI; 2001-589934/66.
DR P-PSDB; AAU28100.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS Claim 1; SEQ ID No 81; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth

CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAS44920-AAS45295 represent novel human secreted protein
CC coding sequences of the invention.
XX
SQ Sequence 2390 BP; 503 A; 710 C; 723 G; 454 T; 0 other;

Query Match 66.2%; Score 19.2; DB 22; Length 2390;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 atctgcagccatggctaggccccc 27
|||||
Db 1958 atctgcagccatggctaggccccc 1981

RESULT 7
AAS77493/C
ID AAS77493 standard; cDNA: 3023 BP.
XX
AC AAS77493;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13297.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG13306.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID NO 13297; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3023 BP; 1083 A; 698 C; 631 G; 611 T; 0 other;

Query Match 66.2%; Score 19.2; DB 23; Length 3023;
Best Local Similarity 87.5%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 tctgcagccatggctagggccccc 28
|||||
Db 1362 TCTGCAGCCATGGCTAGGTGGCG 1339

RESULT 8
AAS80007/c
ID AAS80007 standard; cDNA; 3684 BP.

XX AAS80007;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #15811.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

XX 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG15820.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 1; SEQ ID NO 15811; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 3684 BP; 1391 A; 856 C; 712 G; 725 T; 0 other;

Query Match 66.2%; Score 19.2; DB 23; Length 3684;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 tctgcagccatggctagggccccc 28
|||||
Db 2628 TCTGCAGCCATGGCTAGGTGGCG 2605

RESULT 9
AAL36719

ID AAL36719 standard; DNA; 934 BP.

XX AAL36719;

AC AAL36719;

XX 08-JAN-2002 (first entry)

DT Human musculoskeletal system related polynucleotide SEQ ID NO 3084.

DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.

XX Homo sapiens.

OS WO200155367-A1.

XX 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US01338.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/ or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Example 2; SEQ ID NO 3084; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins,

CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 934 BP; 213 A; 258 C; 241 G; 222 T; 0 other;

Query Match 65.5%; Score 19; DB 22; Length 934;
Best Local Similarity 81.5%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatattctgcagccatggttaggcccc 27
| ||||| | ||||| ||||| |||
Db 66 gctattcttgcctggttaggctccc 92

RESULT 10
AAL36720
ID AAL36720 standard; DNA; 934 BP.
XX
AC AAL36720;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3085.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/ or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Example 2; SEQ ID NO 3085; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the

CC

printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

Sequence 934 BP; 212 A; 258 C; 242 G; 222 T; 0 other;

Query Match 65.5%; Score 19; DB 22; Length 934;
Best Local Similarity 81.5%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatatctgcagccatggttaggcccc 27

| ||||| | ||||| ||||| |||

Db 66 gctatcttctgcatggttaggctccc 92

RESULT 11

AAA37841

ID AAA37841 standard; DNA; 1173 BP.

XX AAA37841;

XX 12-FEB-2001 (first entry)

XX Mouse angiotensin related-2 coding sequence.

AR-2 fibrinogen-like domain; angiotensin related-2; mouse; ischaemia; diabetes; tumour angiogenesis; neoplastic disease; atherosclerosis; thromboembolic disease; inflammatory disease; wound healing; vascularisation; therapy; diagnosis; ds.

OS Mus sp.

XX Key Location/Qualifiers

FT CDS 1..1173

FT /*tag= a

FT /product= angiotensin_related-2

XX WO200052167-A1.

XX 08-SEP-2000.

XX 10-FEB-2000; 2000WO-US03381.

XX 02-MAR-1999; 99US-0122499.

XX (REGE-) REGENERON PHARM INC.

XX Jones PF, Valenzuela DM;

XX WPI; 2000-579286/54.

XX P-PSDB; AAY90401.

Novel nucleic acid molecules encoding fibrinogen-like domain of Angiotensin Related-2 factor useful for treating ischaemia, diabetes, and for wound healing

XX Example 5; Fig 5; 55pp; English.

XX This sequence encodes mouse angiotensin related-2. It was used in the isolation of the angiotensin related-2 (AR-2) fibrinogen-like domain (FD) of the invention. AR-2 is useful for treating and diagnosing ischaemia, diabetes, tumour angiogenesis, neoplastic diseases, thromboembolic diseases, atherosclerosis, inflammatory diseases, and for wound healing. The DNA is useful for developing ligands, screening agonists and antagonists of AR-2, and as a therapeutic for treating disorders involving cells, tissues or organs expressing AR-2 receptor. AR-2 is useful to promote the growth, survival, migration, stabilisation or destabilisation, and/or differentiation of cells expressing AR-2 receptor. AR-2 is also useful in assay systems to identify agonists and antagonists of AR-2 receptor. AR-2 is also useful for inducing or preventing vascularisation in diseases or disorders where such function is indicated, for delivering toxins to a receptor bearing cells and as diagnostic reagents for detecting the disease by tissue staining or

CC whole body imaging.
XX Sequence 1173 BP; 302 A; 334 C; 316 G; 221 T; 0 other;
SQ

Query Match 65.5%; Score 19; DB 21; Length 1173;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 tatctgcagccatggctagggcccccgc 29
|| ||||| || || ||||| ||
Db 311 tacctgcagccagcctatgccccgc 337

RESULT 12
AAS22675
ID AAS22675 standard; cDNA; 1223 BP.
XX
AC AAS22675;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA encoding a novel human protein #241.
XX
KW Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
PN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02623.
XX
PR 25-JAN-2000; 2000US-0491404.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-451939/48.
DR P-PSDB; AAU14370.
XX
PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
PS Claim 1; Page 670-671; 894pp; English.
XX

CC The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/ elicit an immune response, to determine quantitative
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,

CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
XX The present sequence encodes a protein of the invention.
SQ Sequence 1223 BP; 226 A; 357 C; 333 G; 307 T; 0 other;

Query Match 65.5%; Score 19; DB 22; Length 1223;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatattgcagccatggctagggccccc 27
|| | ||||| || ||||| ||
Db 1145 gaaacatcagccatggccagggccccc 1171

RESULT 13
AAS22439
ID AAS22439 standard; cDNA; 1260 BP.
XX
AC AAS22439;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA encoding a novel human protein #5.
XX
KW Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
PN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US02623.
XX
PR 25-JAN-2000; 2000US-0491404.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-451939/48.
DR P-PSDB; AAU14134.
XX
PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
PS Claim 1; Page 172-173; 894pp; English.
XX

CC The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/ elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.

CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence encodes a protein of the invention.
XX
SQ Sequence 1260 BP; 231 A; 371 C; 343 G; 315 T; 0 other;

Query Match 65.5%; Score 19; DB 22; Length 1260;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatattgtcagccatggttagggcccc 27
||| | ||||| ||||| |||||
Db 1182 gaaacatcagccatggttagggcccc 1208

RESULT 14
AAH08668/c
ID AAH08668 standard; cDNA; 702 BP.

XX AC AAH08668;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:5503.

XX DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX PT full-length cDNAs defined in the specification, and for the detection

XX PT and/or diagnosis of the abnormality of the proteins encoded by the

XX PT full-length cDNAs -

XX PS Claim 1; SEQ ID 5503; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesising 5602

XX CC full-length cDNAs defined in the specification. Where a primer set

XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX CC to the complementary strand of a polynucleotide which comprises one of

CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 702 BP; 151 A; 169 C; 243 G; 136 T; 3 other;

Query Match 64.1%; Score 18.6; DB 22; Length 702;
Best Local Similarity 84.0%; Pred. No. 1.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 tctgcagccatggttagggccccgc 29
||| | ||||| ||||| |||||
Db 208 TCTCCGCCCATGGCTCGGCCGCCGC 184

RESULT 15
AAH06512/c

XX ID AAH06512 standard; cDNA; 703 BP.

XX AC AAH06512;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:3347.

XX DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX PT full-length cDNAs defined in the specification, and for the detection

XX PT and/or diagnosis of the abnormality of the proteins encoded by the

XX PT full-length cDNAs -

XX PS Claim 1; SEQ ID 3347; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesising 5602

XX CC full-length cDNAs defined in the specification. Where a primer set

XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX

SQ Sequence 703 BP; 164 A; 163 C; 232 G; 141 T; 3 other;

Query Match 64.1%; Score 18.6; DB 22; Length 703;
 Best Local Similarity 84.0%; Pred. No. 1.8e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 tctgcagccatggctagggcccccgc 29
 |||||
 Db 150 TCTCCCGCCATGGCTCGGCGGCCGC 126

Search completed: August 31, 2002, 18:17:15
 Job time: 16437 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 13:38:17 ; Search time 6064.26 seconds
(without alignments)
64.544 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattgcagccatggctagggccccgc 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	72.4	901	12	CNS03EI9
2	20.8	71.7	365	10	BG697507
3	20.2	69.7	687	12	AG125785
4	20	69.0	421	12	AQ458858
5	20	69.0	1700	10	BG111410
6	19.6	67.6	383	9	AW346670
7	19.6	67.6	554	9	AW657502
8	19.6	67.6	563	10	BG383058
9	19.4	66.9	461	10	BE356665
10	19.4	66.9	574	9	AW066213
11	19.4	66.9	1249	10	BF690022
12	19.2	66.2	240	10	BF516864
13	19.2	66.2	306	9	AW405340
14	19.2	66.2	315	9	AW736930
15	19.2	66.2	326	10	BE761821
16	19.2	66.2	330	9	AW056925
17	19.2	66.2	330	10	BE582369

18	19.2	66.2	348	10	BF220344
19	19.2	66.2	400	10	BF777295
20	19.2	66.2	423	10	H02247
21	19.2	66.2	443	10	BE761860
22	19.2	66.2	452	10	BE656912
23	19.2	66.2	518	12	AZ827981
24	19.2	66.2	519	10	BG040220
25	19.2	66.2	534	10	BF610423
26	19.2	66.2	538	10	BF610478
27	19.2	66.2	548	9	AW290141
28	19.2	66.2	591	9	AW064983
29	19.2	66.2	618	11	AK017620
30	19.2	66.2	983	9	AL572399
31	19.2	66.2	1100	10	BG286946
32	19	65.5	388	12	AZ995875
33	19	65.5	452	12	AQ495929
34	19	65.5	490	10	BJ202596
35	19	65.5	554	10	BJ205326
36	19	65.5	561	10	BJ202480
37	19	65.5	563	10	BJ192283
38	19	65.5	573	10	BJ156941
39	19	65.5	579	10	BJ200115
40	19	65.5	594	10	BF282124
41	19	65.5	673	12	AZ359142
42	19	65.5	1125	10	BG025522
43	19	65.5	2306	11	AK011976
44	18.8	64.8	150	10	BF604299
45	18.8	64.8	227	9	AA556708

ALIGNMENTS

RESULT 1
CNS03EI9
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
020G12 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL240426
VERSION
AL240426.1 GI:7961195
KEYWORDS
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis.
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 901)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 901)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL
Unpublished
REFERENCE
3 (bases 1 to 901)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
JOURNAL
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Location/Qualifiers
1..901
/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"
/clone="020G12"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG020BD06SP1-end :
PUC-ori"
BASE COUNT 254 a 239 c 228 g 173 t 7 others
ORIGIN

Query Match 72.4%; Score 21; DB 12; Length 901;
Best Local Similarity 82.8%; Pred. No. 4.5e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatatctgcagccatggctaggccccgc 29
|||||
Db 154 GATCTCTGCAGCCATCGCTCGGCTCCCC 182

RESULT 2
BG697507/c
LOCUS 602660715F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4804011 5',
DEFINITION mRNA sequence.
ACCESSION BG697507
VERSION BG697507.1 GI:13963799
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 365)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10700 row: m column: 04
High quality sequence start: 9
High quality sequence stop: 365.
FEATURES
source
1..365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4804011"
/clone_lib="NCI_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: Sali; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 62 a 117 c 121 g 65 t
ORIGIN

Query Match 71.7%; Score 20.8; DB 10; Length 365;
Best Local Similarity 91.7%; Pred. No. 4.4e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ctgcagccatggctaggccccgc 29
|||||
Db 54 CTGCAGCCATGGCCAGGCCCCAGC 31

RESULT 3
AG125785/c
LOCUS 687 bp DNA linear GSS 04-NOV-2001

DEFINITION Pan troglodytes DNA, clone: PTB-136C09.F, genomic survey sequence.
ACCESSION AG125785
VERSION AG125785.1 GI:16654950
KEYWORDS GSS; GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC Library clone:PTB-136C09.F.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (sites)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 687)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpbes@gsc.riken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
source
1..687
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-136C09.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 190 a 142 c 140 g 213 t 2 others
ORIGIN

Query Match 69.7%; Score 20.2; DB 12; Length 687;
Best Local Similarity 88.0%; Pred. No. 8.5e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gatatctgcagccatggctaggccc 25
|||||
Db 331 GATATCTGAAGCCCATGCCAAGACCC 307

RESULT 4
AQ458858
LOCUS HS_5050_A1_A10_SP6E RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate=626 Col=19 Row=A, DNA sequence.
ACCESSION AQ458858
VERSION AQ458858.1 GI:4637498
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 421)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@defong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 626 row: A column: 19
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 421.

FEATURES

source
Location/Qualifiers
1. .421
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-626 Col-19 Row-A"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"

BASE COUNT 136 a 70 c 69 g 145 t 1 others
ORIGIN

Query Match 69.0%; Score 20; DB 12; Length 421;
Best Local Similarity 82.1%; Pred. No. 9.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 atatctgcagccatggctaggccccgc 29
||||||| ||| |||| |||| ||
Db 251 ATATCTGCAGGCATTGTTAGGACCCAGC 278

RESULT 5
BG111410/c

LOCUS
DEFINITION BG111410 1700 bp mRNA linear EST 30-JAN-2001
602881779F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4369363 5',
mRNA sequence.

ACCESSION BG111410
VERSION BG111410.1 GI:12604832
KEYWORDS EST.
SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1700)

NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10024 row: n column: 20

High quality sequence stop: 290.
Location/Qualifiers
1. .1700

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:4369363"
/clone_lib="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.533 kb. Library enriched for full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 582 a 552 c 451 g 115 t
ORIGIN

Query Match 69.0%; Score 20; DB 10; Length 1700;
Best Local Similarity 82.1%; Pred. No. 1.2e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatatctgcagccatggctaggccccg 28
||||| ||||| ||||| ||||| ||
Db 594 GGTATCTCCGCCATGGCTAGGCTCGCG 567

RESULT 6
AW346670

LOCUS
DEFINITION AW346670 383 bp mRNA linear EST 09-JUL-2000
29219 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AW346670
VERSION AW346670.1 GI:6844380
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

REFERENCE

1 (bases 1 to 383)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.

TITLE

Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL

JOURNAL

COMMENT

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 13 row: P column: 6
Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

Location/Qualifiers
1. .383
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

BASE COUNT 89 a 141 c 103 g 50 t
ORIGIN

Query Match 67.6%; Score 19.6; DB 9; Length 383;
Best Local Similarity 84.6%; Pred. No. 1.3e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 atatctgcagccatggctaggcccc 27

Db 346 ATACCTGCAGCCATGGCATGGGCCCC 371
RESULT 7
AW657502
LOCUS AW657502 554 bp mRNA linear EST 09-JUL-2000
DEFINITION 110747 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AW657502
VERSION AW657502.1 GI:7423328
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 554)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
JOURNAL EST discovery in swine
COMMENT Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 38 row: I column: 2
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1..554
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 117 a 198 c 154 g 85 t
ORIGIN
Query Match 67.6%; Score 19.6; DB 9; Length 554;
Best Local Similarity 84.6%; Pred. No. 1.4e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 atatctgcagccatggcctaggcccc 27
||| ||||| ||||| ||| |||||
Db 419 ATACCTGCAGCCATGGCATGGGCCCC 444
RESULT 8
BG383058/c
LOCUS BG383058 563 bp mRNA linear EST 12-MAR-2001
DEFINITION 300957 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG383058
VERSION BG383058.1 GI:13307530
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 563)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.

and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 6 row: C column: 20
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1..563
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 165 a 148 c 129 g 120 t 1 others
ORIGIN
Query Match 67.6%; Score 19.6; DB 10; Length 563;
Best Local Similarity 84.6%; Pred. No. 1.4e+03;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 atatctgcagccatggcctaggcccc 27
||| ||||| ||||| ||||| |||||
Db 359 ATGAATGCAGCCCAAGGCTAGGCCCCC 334
RESULT 9
BE356665/c
LOCUS BE356665 461 bp mRNA linear EST 20-JUL-2000
DEFINITION DGL_12_A12_b1_A002 Dark Grown 1 (DGL) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION BE356665
VERSION BE356665.1 GI:9298222
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 461)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude polyA, vector and regions
below phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 451
POLYA-No.
FEATURES Location/Qualifiers

```
source
1. .461
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGL)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT      67 a   160 c   154 g   80 t
ORIGIN

Query Match      66.9%; Score 19.4; DB 10; Length 461;
Best Local Similarity 79.3%; Pred. No. 1.6e+03;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 gatattgcagccatggctagggcccccgc 29
||||| ||||| ||||| ||||| |||||
Db 373 GATGTCAGCAGCGAGGCGGAGGTCCCGC 345

RESULT 10
AW066213/c
LOCUS
DEFINITION
AW066213 574 bp mRNA linear EST 30-MAR-2000
687008A06.y1 687 - Early embryo from Delaware Zea mays cDNA, mRNA
sequence.
ACCESSION
AW066213
VERSION
AW066213.1 GI:6021285
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 574)
AUTHORS
Walbot,V.
TITLE
Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL
Unpublished (1999)
COMMENT
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 687008 row: A column: 06.
Location/Qualifiers
1. .574
/organism="Zea mays"
/cultivar="Illinois High Oil"
/db_xref="taxon:4577"
/clone_lib="687 - Early embryo from Delaware"
/tissue_type="embryo"
/dev_stage="14, 21, 28, and 35 days after pollination"
/lab_host="E. coli SOLR"
/note="Organ: embryo; Vector: pBluescript SK; Site_1: XhoI
; Site_2: EcoRI; Library was prepared by Statagene using
the Uni-ZAP XR system (Stratagene BN937328-12). Clones
were picked by a Q-bot after blue/white selection
(ampicillin resistance - use 100 micrograms/microliter).
Developed from a pool of equal amounts of RNA from
developing embryos sampled at 14, 21, 28 and 35 days after
pollination of the Illinois High Oil Maize Strain Cycle
90. This closed strain has been selected for high oil
concentration for 90 generations and originates from the
1890s era open pollinated variety Burr's White"
BASE COUNT      101 a   181 c   183 g   108 t   1 others
ORIGIN

Query Match      66.9%; Score 19.4; DB 9; Length 574;
Best Local Similarity 79.3%; Pred. No. 1.6e+03;
```

```
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 gatattgcagccatggctagggcccccgc 29
||||| ||||| ||||| ||||| |||||
Db 402 GATATCAGCGCGAGGCGGAGGTCCCGC 374

RESULT 11
BF690022/c
LOCUS
DEFINITION
BF690022 1249 bp mRNA linear EST 22-DEC-2000
602186312T1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4298522 3',
mRNA sequence.
ACCESSION
BF690022
VERSION
BF690022.1 GI:11975430
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1249)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1154 row: g column: 03
High quality sequence start: 30
High quality sequence stop: 509.
Location/Qualifiers
1. .1249
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4298522"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      336 a   303 c   368 g   241 t   1 others
ORIGIN

Query Match      66.9%; Score 19.4; DB 10; Length 1249;
Best Local Similarity 79.3%; Pred. No. 1.9e+03;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 gatattgcagccatggctagggcccccgc 29
||||| ||||| ||||| ||||| |||||
Db 60 GATATCCGACGCCGTGGCTAGGATCCAC 32

RESULT 12
BF516864
LOCUS
DEFINITION
BF516864 240 bp mRNA linear EST 08-DEC-2000
NXSI_004_D07_F NXSI (Nsf Xylem Side wood Inclined) Pinus taeda cDNA
clone NXSI_004_D07 5', mRNA sequence.
ACCESSION
BF516864
VERSION
BF516864.1 GI:11604145
KEYWORDS
EST.
```

SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 240)
Sederoff, R.
Molecular Basis of Wood Formation in the Pine Megagenome
Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: ajohnson@unity.ncsu.edu
Seq primer: T3.

FEATURES
source
1..240
/organism="pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXSI_004_D07"
/clone_lib="NXSI (Nsf xylem Side wood Inclined)"
/tissue_type="Xylem"
/cell_type="Side"
/dev_stage="Juvenile"
/lab_host="XL1-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form side
wood by bending to a 45 degree angle and tying them to the
ground. Differentiating xylem was harvested from the sides
of the inclined stems, and a mixture of all three
genotypes was used for the library. oligo-dT primed cdna
was directionally cloned into the EcoRI-XhoI Bluescript SK
vector arms. NOTE: The sequences contain a 'cdna adapter'
between the EcoRI site and the start of the EST. The
adapter sequence is 'AATTCGGCAGCAG'."

BASE COUNT 55 a 70 c 63 g 44 t 8 others
ORIGIN
Query Match 66.2%; Score 19.2; DB 10; Length 240;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 tctgcagccatggctaggcccccg 28
||||| ||||| || |||||
Db 27 TCTGAAGCCATGGCAAGACCCCG 50

RESULT 13
AW405340
LOCUS
DEFINITION
UI-HF-BL0-ack-f-06-0-UI.r1 NIH_MGC_37 Homo sapiens cdna clone
IMAGE:3059459 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 306)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward.

FEATURES
source
1..306
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3059459"
/clone_lib="NIH_MGC_37"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/note="Vector: pT7T3-pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(1.5-2.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 61 a 117 c 88 g 40 t
ORIGIN

Query Match 66.2%; Score 19.2; DB 9; Length 306;
Best Local Similarity 87.5%; Pred. No. 1.7e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 atctgcagccatggctaggcccc 27
||||| ||||| || |||||
Db 272 ATTCGAGCCATGGCATGGACCCC 295

RESULT 14
AW736930
LOCUS
DEFINITION
NXNV_081_F12_F Nsf Xylem Normal wood Vertical Pinus taeda cdna
clone NXNV_081_F12 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
AW736930.1 GI:7644794
EST.
loblolly pine.
ORGANISM
Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 315)
Sederoff, R.
Molecular Basis of Wood Formation in the Pine Megagenome
Unpublished (2000)
Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: ajohnson@unity.ncsu.edu
Seq primer: T3.

FEATURES
source
1..315
Location/Qualifiers
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="NXNV_081_F12"
/clone_lib="Nsf Xylem Normal wood Vertical"
/note="Vector: Bluescript SK; Site_1: Eco RI; The
sequences contain a 'cdna adapter' between the EcoRI site
and the start of the EST. The adapter sequence is
'AATTCGGCAGCAG'."

BASE COUNT 79 a 87 c 68 g 70 t 11 others
ORIGIN

Query Match 66.2%; Score 19.2; DB 9; Length 315;
Best Local Similarity 87.5%; Pred. No. 1.7e+03;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 tctgcagccatggctaggcccccg 28

Db155TCTGAAGCCCATGGCAAGACCCCG178|||||||||||||

RESULT15

BE761821

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

OY

Db

Search completed: August 31, 2002, 15:24:33
Job time: 6376 sec

THIS PAGE BLANK (Ver. 3.0)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 13:40:38 ; Search time 153.1 Seconds
(without alignments)
46.528 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattgcagccatggttaggcgccgcg 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	18.2	62.8	3107	1 US-07-846-181-4	Sequence 4, Appl1
C 2	18.2	62.8	3107	1 US-07-845-989-4	Sequence 4, Appl1
C 3	17.8	61.4	1774	1 US-08-377-292-3	Sequence 3, Appl1
C 4	17.8	61.4	4353	2 US-08-365-486A-18	Sequence 18, Appl1
C 5	17.8	61.4	4353	4 US-08-880-342-18	Sequence 18, Appl1
C 6	17.8	61.4	4780	2 US-08-365-486A-20	Sequence 20, Appl1
C 7	17.8	61.4	4780	3 US-09-123-708-3	Sequence 3, Appl1
C 8	17.8	61.4	4780	3 US-09-123-624-3	Sequence 3, Appl1
C 9	17.8	61.4	4780	4 US-08-880-342-20	Sequence 20, Appl1
C 10	17.6	60.7	652	4 US-08-998-416-962	Sequence 962, App
C 11	17.2	59.3	31	2 US-08-890-980-80	Sequence 80, Appl1
C 12	17.2	59.3	31	2 US-08-890-980-82	Sequence 82, Appl1
C 13	17.2	59.3	31	3 US-08-890-979-71	Sequence 71, Appl1
C 14	17.2	59.3	31	3 US-08-890-979-73	Sequence 73, Appl1
C 15	17.2	59.3	31	3 US-09-032-894-80	Sequence 80, Appl1
C 16	17.2	59.3	31	3 US-09-032-894-82	Sequence 82, Appl1
C 17	17.2	59.3	31	4 US-09-031-626-80	Sequence 80, Appl1
C 18	17.2	59.3	31	4 US-09-031-626-82	Sequence 82, Appl1
C 19	17.2	59.3	160	2 US-08-890-980-66	Sequence 66, Appl1
C 20	17.2	59.3	160	3 US-08-890-979-65	Sequence 65, Appl1
C 21	17.2	59.3	160	3 US-09-032-894-66	Sequence 66, Appl1
C 22	17.2	59.3	160	4 US-09-031-626-66	Sequence 66, Appl1
C 23	17	58.6	730	4 US-09-342-084-5	Sequence 5, Appl1
C 24	16.8	57.9	800	4 US-08-998-416-390	Sequence 390, App
C 25	16.8	57.9	5975	1 US-08-920-812-23	Sequence 23, Appl1
C 26	16.8	57.9	5975	1 US-08-920-827-23	Sequence 23, Appl1
C 27	16.8	57.9	5975	1 US-08-921-177-23	Sequence 23, Appl1

28	16.8	57.9	5975	1	US-08-362-577C-23	Sequence 23, Appl1
29	16.8	57.9	5975	2	US-08-920-828-23	Sequence 23, Appl1
C 30	16.6	57.2	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
C 31	16.6	57.2	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
C 32	16.4	56.6	1704	1	US-08-007-282B-1	Sequence 1, Appl1
C 33	16.2	55.9	376	1	US-08-785-530-2	Sequence 2, Appl1
C 34	16.2	55.9	376	2	US-09-123-850-2	Sequence 2, Appl1
C 35	16.2	55.9	515	4	US-09-060-756-656	Sequence 656, App
C 36	16.2	55.9	638	4	US-08-858-207A-142	Sequence 142, App
C 37	16.2	55.9	779	4	US-08-998-416-451	Sequence 451, App
C 38	16.2	55.9	1685	1	US-07-708-038-1	Sequence 1, Appl1
C 39	16.2	55.9	1685	1	US-08-127-995-1	Sequence 1, Appl1
C 40	16.2	55.9	5057	2	US-08-365-486A-12	Sequence 12, Appl1
C 41	16.2	55.9	5057	4	US-08-880-342-12	Sequence 12, Appl1
C 42	16.2	55.9	5108	1	US-07-642-002-1	Sequence 1, Appl1
C 43	16.2	55.9	9143	2	US-08-639-857-32	Sequence 32, Appl1
C 44	16.2	55.9	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
C 45	16.2	55.9	4411529	4	US-09-103-840A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-07-846-181-4/c
; Sequence 4, Application US/07846181
; Patent No. 5360732
; GENERAL INFORMATION:
; APPLICANT: BERKA, RANDY M
; APPLICANT: FOWLER, TIMOTHY
; APPLICANT: REY, MICHAEL W
; TITLE OF INVENTION: PRODUCTION OF ASPERGILLUS NIGER
; TITLE OF INVENTION: CATALASE-R
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 180 KIMBALL WAY
; CITY: SOUTH SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/846,181
; FILING DATE: 19920304
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HORN Ms, MARGARET A
; REGISTRATION NUMBER: 33401
; REFERENCE/DOCKET NUMBER: GC204-US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-742-7536
; TELEFAX: 415-742-7217
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3107 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(327..620, 683..907, 969..1385, 1440..1604,
; LOCATION: 1654..2745)
US-07-846-181-4

Query Match 62.8%; Score 18.2; DB 1; Length 3107;
Best Local Similarity 87.0%; Pred. No. 51;
.Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 atatctgcagccatggctaggcc 24
||||| ||| ||||| |||||
Db 242 ATATCGGCATCCATGGCGAGGCC 220

RESULT 2
US-07-845-989-4/c
; Sequence 4, Application US/07845989
; Patent No. 5360901
; GENERAL INFORMATION:
; APPLICANT: BERKA, RANDY M
; APPLICANT: FOWLER, TIMOTHY
; APPLICANT: REY, MICHAEL W
; TITLE OF INVENTION: PRODUCTION OF ASPERGILLUS NIGER
; TITLE OF INVENTION: CATALASE-R
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.
; STREET: 180 KIMBALL WAY
; CITY: SOUTH SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/845,989
; FILING DATE: 19920304
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HORN MS, MARGARET A
; REGISTRATION NUMBER: 33401
; REFERENCE/DOCKET NUMBER: GC208-US1
; TELEPHONE: 415-742-7536
; TELEFAX: 415-742-7217
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3107 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; NAME/KEY: CDS
; LOCATION: Join(327..620, 683..907, 969..1385, 1440..1604,
; LOCATION: 1654..2745)
; US-07-845-989-4

Query Match 62.8%; Score 18.2; DB 1; Length 3107;
Best Local Similarity 87.0%; Pred. No. 51;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 atatctgcagccatggctaggcc 24
||||| ||| ||||| |||||
Db 242 ATATCGGCATCCATGGCGAGGCC 220

RESULT 3
US-08-377-292-3/c
; Sequence 3, Application US/08377292
; Patent No. 5693615
; GENERAL INFORMATION:
; APPLICANT: STONE, ROGER L.
; TITLE OF INVENTION: THERAPEUTIC FORMULAS FOR OSTEOINDUCTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road

; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 45239-8707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,292
; FILING DATE: 23-JAN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,435
; FILING DATE:
; APPLICATION NUMBER: US/08/117,367
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Corstanje, Brahm J.
; REGISTRATION NUMBER: 34,804
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-245-2858
; TELEFAX: 513-741-3012
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-377-292-3

Query Match 61.4%; Score 17.8; DB 1; Length 1774;
Best Local Similarity 90.5%; Pred. No. 171;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 cagccatggctaggcccccgc 29
||||| ||||| |||||
Db 327 CAGCCATGGCTAGGTACCCGC 307

RESULT 4
US-08-365-486A-18/c
; Sequence 18, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/365,486A
; FILING DATE: 23-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8255-0018

; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "probe"
US-08-890-980-82

Query Match 59.3%; Score 17.2; DB 2; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 gcagccatggctagggcccccgc 29
|||||
Db 28 GCAGCCATGGCCAGGCCACCC 7

RESULT 13
US-08-890-979-71
; Sequence 71, Application US/08890979
; Patent No. 6030778
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Ordovas, Jose M.
; TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,979
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-005.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "probe"
US-08-890-979-71

Query Match 59.3%; Score 17.2; DB 3; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 gcagccatggctagggcccccgc 29
|||||
Db 4 GCAGCCATGGCCAGGCCACCC 25

RESULT 14
US-08-890-979-73/c
; Sequence 73, Application US/08890979
; Patent No. 6030778
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; APPLICANT: Ordovas, Jose M.
; TITLE OF INVENTION: DIAGNOSTIC ASSAYS AND KITS FOR BODY MASS
; TITLE OF INVENTION: DISORDERS
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,979
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MIA-005.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "probe"
US-08-890-979-73

Query Match 59.3%; Score 17.2; DB 3; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 gcagccatggctagggcccccgc 29
|||||
Db 28 GCAGCCATGGCCAGGCCACCC 7

RESULT 15
US-09-032-894-80
; Sequence 80, Application US/09032894
; Patent No. 6130041
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: SR-BI NUCLEIC ACIDS AND USES THEREFOR
; FILE REFERENCE: MIA-005.03
; CURRENT APPLICATION NUMBER: US/09/032,894
; CURRENT FILING DATE: 1998-02-27
; EARLIER APPLICATION NUMBER: 08/890,980
; EARLIER FILING DATE: 1997-07-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Human

US-09-032-894-80

Query Match 59.3%; Score 17.2; DB 3; Length 31;
Best Local Similarity 86.4%; Pred. No. 85;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 gcagccatggctagggcccccgc 29
| | | | | | | | | | | | | | | | | |
Db 4 gcagccatggccagggccacc 25

Search completed: August 31, 2002, 15:27:36
Job time: 6418 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:03:50 ; Search time 9324.87 Seconds
(without alignments)
69.569 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtagcctacaggtagcgtgagcaatttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Match	Length	ID Description

1	31	100.0	31	6	AX275253	Sequence
c	31	100.0	5767	6	AX275254	Sequence
c	31	100.0	14446	6	AX275255	Sequence
c	20.4	65.8	1725	6	AX275256	Sequence
c	20.4	65.8	136257	9	AC073864	Homo sapi
c	20.4	65.8	170075	2	AC026678	Homo sapi
c	20.4	65.8	172241	2	AC024508	Homo sapi
c	20.4	65.8	178820	9	AC020552	Homo sapi
c	20.4	65.8	181230	2	AL358876	Homo sapi
c	20.4	65.8	183754	2	AC023183	Homo sapi
c	20.4	65.8	260430	2	AC024935	Homo sapi
c	20.4	64.5	1019	1	DSP276564	Desulfoto
c	20.4	64.5	1409	1	AB015893	Treponema
c	20.4	64.5	1410	1	AB015888	Treponema
c	20.4	64.5	1414	1	AB015889	Treponema
c	20.4	64.5	1417	1	AB015887	Treponema
c	20.4	64.5	1421	1	AB015892	Treponema
c	20.4	64.5	1475	1	AF033303	Treponema
c	20.4	64.5	1479	1	AF033307	Treponema
c	20.4	64.5	1543	1	AF033306	Treponema
c	20.4	64.5	87600	9	AC007160	Homo sapi
c	19.8	63.9	478	1	USP287095	unculture
c	19.8	63.9	1469	1	AF023038	Treponema
c	19.8	63.9	32147	9	AL358613	Human DNA
c	19.8	63.9	348550	1	AP003596	Nostoc sp
c	19.6	63.2	73814	2	AC069345	Homo sapi
c	19.6	63.2	95180	9	AC098853	Homo sapi
c	19.6	63.2	218677	2	AC104842	Homo sapi
c	19.4	62.6	461	1	USP287095	unculture
c	19.4	62.6	91316	2	AC094771	Rattus no
c	19.4	62.6	93957	2	AC021819	Homo sapi
c	19.4	62.6	110000	2	LMFLCHR36_27	Continuation (28 o
c	19.4	62.6	110000	2	LMFLCHR36_28	Continuation (29 o
c	19.4	62.6	141180	2	AC078807	Homo sapi
c	19.4	62.6	141203	2	AC068635	Homo sapi
c	19.4	62.6	170940	2	AC069067	Homo sapi
c	19.4	62.6	180681	9	AC017016	Homo sapi
c	19.2	61.9	1471	1	BLSFO16SN	X80834 Long segmen
c	19	61.3	1637	1	AF231923	Xanthomon
c	19	61.3	64414	9	AC023792	Homo sapi
c	19	61.3	145540	2	AC015962	Homo sapi
c	19	61.3	171718	2	AC022501	Homo sapi
c	19	61.3	173037	2	AC018542	Homo sapi
c	19	61.3	173487	2	AC026925	Homo sapi
c	19	61.3	174803	9	AL591501	Human DNA

ALIGNMENTS

RESULT	1				
AX275253	AX275253	Sequence 2 from Patent WO0171014.	31 bp	DNA	linear PAT 29-OCT-2001
LOCUS	AX275253	Sequence 2 from Patent WO0171014.			
DEFINITION	AX275253				
ACCESSION	AX275253				
VERSION	AX275253.1	GI:16547673			
KEYWORDS					
SOURCE		synthetic construct.			
ORGANISM		synthetic construct			
REFERENCE		artificial sequence.			
AUTHORS		1 (sites)			
TITLE		Mor,T., Soreq,H., Arntzen,C. and Mason,H.			
JOURNAL		Expression of recombinant human acetylcholinesterase in transgenic plants			
Patent:		WO 0171014-A 2 27-SEP-2001;			
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,					
Tsaafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,					
Hugh S. (US)					
FEATURES		Location/Qualifiers			
source		1..31			
		/organism="synthetic construct"			
		/db_xref="taxon:32630"			
		/note="primer pAcHE-Kpn, derived from human AcHE gene and			

BASE COUNT 7 a 7 c 9 g 8 t
ORIGIN modified to introduce a Kpn I restriction site"

Query Match 100.0%; Score 31; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cggtagctatcaggtagcgctgagcaatttg 31
|||||
Db 1 CGGTACCTATCAGGTAGCGCTGAGCAATTG 31

RESULT 2
AX275254/c
LOCUS AX275254 5767 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 3 from Patent WO0171014.
ACCESSION AX275254
VERSION AX275254.1 GI:16547674
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE
AUTHORS 1 (sites)
Mor, T., Soreq, H., Arntzen, C. and Mason, H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants
JOURNAL Patent: WO 0171014-A 3 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsaafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES
source
1..5767
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM034"

BASE COUNT 1390 a 1521 c 1495 g 1361 t
ORIGIN

Query Match 100.0%; Score 31; DB 6; Length 5767;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cggtagctatcaggtagcgctgagcaatttg 31
|||||
Db 2571 CGGTACCTATCAGGTAGCGCTGAGCAATTG 2541

RESULT 3
AX275255/c
LOCUS AX275255 14446 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 4 from Patent WO0171014.
ACCESSION AX275255
VERSION AX275255.1 GI:16547675
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE
AUTHORS 1 (sites)
Mor, T., Soreq, H., Arntzen, C. and Mason, H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants
JOURNAL Patent: WO 0171014-A 4 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsaafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES
source
1..14446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM036"

BASE COUNT 3231 a 3831 c 4046 g 3042 t 296 others
ORIGIN

Query Match 100.0%; Score 31; DB 6; Length 14446;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cggtagctatcaggtagcgctgagcaatttg 31
|||||
Db 1875 CGGTACCTATCAGGTAGCGCTGAGCAATTG 1845

RESULT 4
AX275256/c
LOCUS AX275256 1725 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 5 from Patent WO0171014.
ACCESSION AX275256
VERSION AX275256.1 GI:16547676
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE
AUTHORS 1 (sites)
Mor, T., Soreq, H., Arntzen, C. and Mason, H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants
JOURNAL Patent: WO 0171014-A 5 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsaafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES
source
1..1725
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="synthetic human acetylcholinesterase gene optimized for expression in plants"

BASE COUNT 273 a 553 c 571 g 328 t
ORIGIN

Query Match 65.8%; Score 20.4; DB 6; Length 1725;
Best Local Similarity 95.5%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 tcaggtagcgctgagcaatttg 31
|||||
Db 1725 TCAGGTGGCGCTGAGCAATTG 1704

RESULT 5
AC073864/c
LOCUS AC073864 136257 bp DNA linear PRI 30-MAR-2001
DEFINITION Homo sapiens 12q BAC RP11-227B21 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC073864
VERSION AC073864.28 GI:13489126
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS 1 (bases 1 to 136257)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Banks, T., Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,

Dugan-Rocha,S., Durbin,K.J., Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoostari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 136257)
Worley,K.C.

Direct Submission
Submitted (30-JUN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 136257)
Worley,K.C.

Direct Submission
Submitted (30-MAR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 30, 2001 this sequence version replaced g1:13443043.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-helpebcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for

a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

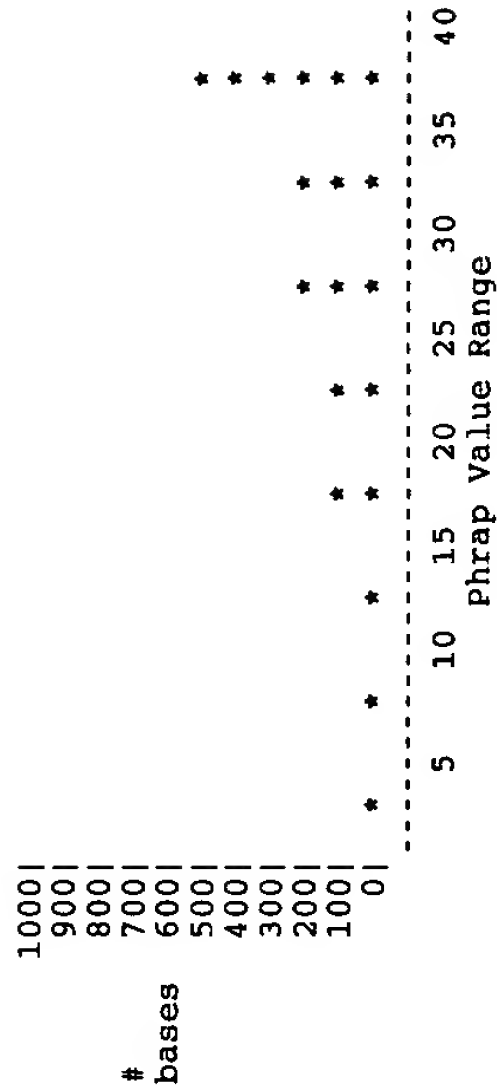
QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 136257
Phrap values in estimate: 135295
Average error rate (BCM-Phrap estimate): 0.00010428
Fraction of Phrap values less than 40 : 0.0103625
Number of consensus changing edits: 20
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
2846 tgataataa(n)caacaataa tgataataa(a)caacaataa
3777 agataaattat(n)atnttgttct agataaattat(t)attttgttct
3780 taattatnat(n)ttgttcttct taattatnat(t)ttgttcttct
17066 caaaaaaaca(n)acagatacat caaaaaaaca(t)acagatacat
17341 tgacagtggg(n)caagtcattt tgacagtggg(t)caagtcattt
18186 atcaaaacttt(n)ctatgaacac atcaaaacttt(a)ctatgaacac
85529 gtgtgcatgt(n)agggtgtgtg gtgtgcatgt(g)agggtgtgtg
85548 tgtgtgtgta(n)gnngngtctg tgtgtgtgta(t)ggagtgtgtc
85550 tgtgtgtang(n)ngngngtctg tgtgtgtang(g)agggtgtgtg
85551 gtgtgtangn(n)ngngtctgga gtgtgtangg(a)gtgtgtctga
85553 gtgtangng(n)ngtctgagg gtgtangngg(t)gtctgagggt
85555 gtangnnng(n)gtctgaggn gtangnnngg(t)gtctgagggt
85564 gngtctgagg(n)ngtngtggag gngtctgagg(g)gtgtgtgga
85565 ngctgaggn(n)gtngtggaga ngctgaggn(g)gtngtggaga
85569 tgaggnngt(n)gtgagaagat tgaggnngt(g)gtgagaagat
93974 ttttttttg(g)ggagacaagg ttttttttg(t)ggagacaagg
107496 aaaaaaaaa(n)caaaaacct aaaaaaaaa(a)caaaaacct
122522 tccttgatt(n)tattatacaa tccttgatt(t)tattatacaa
122608 tggcctaggg(n)ctgttttcc tggcctaggg(c)ctgttttcc
134275 acatgcatgt(n)tcaatcacat acatgcatgt(t)tcaatcacat

----- Distribution of Quality < 40 Bases -----



Version: 1.01 qxfo.
FEATURES Location/Qualifiers
source 1. .136257
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12q"
/clone="RP11-227B21"
repeat_region complement(256..742)

```
source
1. .170075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-220C5"
BASE COUNT 45458 a 38815 c 38968 g 45324 t 1510 others
ORIGIN

Query Match 65.8%; Score 20.4; DB 2; Length 170075;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtagcgtgagcaatttg 31
|||||  ||||  ||||  ||||  ||||
Db 67708 GGTACCTATCTCGTAGAGCTGATAGATTG 67737

RESULT 7
AC024508/c
LOCUS
DEFINITION
Homo sapiens chromosome 1 clone RP11-156A20 map 1, WORKING DRAFT
SEQUENCE, 28 unordered pieces.
AC024508
AC024508.4 GI:10047852
VERSION
KEYWORDS
SOURCE
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 172241)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 1, clone RP11-156A20
Unpublished
2 (bases 1 to 172241)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavkiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe1,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczky,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., Mcpheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivar,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Travers,M., Trigilio,J., Vassiliev,H., Vlel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 10, 2000 this sequence version replaced gi:8099049.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7204
Center clone name: 156_A_20
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
```

```
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 155372 bases at least Q40
Consensus quality: 163469 bases at least Q30
Consensus quality: 166829 bases at least Q20
Insert size: 157000; agarose-fp
Quality coverage: 4.3 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
1 779: contig of 779 bp in length
780 879: gap of 100 bp
880 2388: contig of 1509 bp in length
2389 2488: gap of 100 bp
2489 4256: contig of 1768 bp in length
4257 4356: gap of 100 bp
4357 5784: contig of 1428 bp in length
5785 5884: gap of 100 bp
5885 7892: contig of 2008 bp in length
7893 7992: gap of 100 bp
7993 9655: contig of 1663 bp in length
9656 9755: gap of 100 bp
9756 12205: contig of 2450 bp in length
12206 12305: gap of 100 bp
12306 14310: contig of 2005 bp in length
14311 14410: gap of 100 bp
14411 16707: contig of 2297 bp in length
16708 16807: gap of 100 bp
16808 19279: contig of 2472 bp in length
19280 19379: gap of 100 bp
19380 21937: contig of 2558 bp in length
21938 22037: gap of 100 bp
22038 48726: contig of 26689 bp in length
48727 48826: gap of 100 bp
48827 51847: contig of 3021 bp in length
51848 51947: gap of 100 bp
51948 54626: contig of 2679 bp in length
54627 54726: gap of 100 bp
54727 58944: contig of 4218 bp in length
58945 59044: gap of 100 bp
59045 63316: contig of 4272 bp in length
63317 63416: gap of 100 bp
63417 68257: contig of 4841 bp in length
68258 68357: gap of 100 bp
68358 73280: contig of 4923 bp in length
73281 73380: gap of 100 bp
73381 78209: contig of 4829 bp in length
78210 78309: gap of 100 bp
78310 84542: contig of 6233 bp in length
84543 84642: gap of 100 bp
84643 89899: contig of 5257 bp in length
89900 89999: gap of 100 bp
90000 96909: contig of 6910 bp in length
96910 97009: gap of 100 bp
97010 105708: contig of 8699 bp in length
105709 105808: gap of 100 bp
105809 116467: contig of 10659 bp in length
116468 116567: gap of 100 bp
116568 128407: contig of 11840 bp in length
128408 128507: gap of 100 bp
128508 144528: contig of 16021 bp in length
144529 144628: gap of 100 bp
144629 169246: contig of 24618 bp in length
169247 169346: gap of 100 bp
169347 172241: contig of 2895 bp in length.
```

FEATURES		Location/Qualifiers			
source		1. .172241			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/chromosome="1"			
		/map="1"			
		/clone="RP11-156A20"			
		/clone_lib="RPC1-11 Human Male BAC"			
misc_feature		1. .779			
		/note="assembly_fragment"			
		clone_end:SP6			
		vector_side:left"			
misc_feature		880. .2388			
		/note="assembly_fragment"			
misc_feature		2489. .4256			
		/note="assembly_fragment"			
misc_feature		4357. .5784			
		/note="assembly_fragment"			
misc_feature		5885. .7892			
		/note="assembly_fragment"			
misc_feature		7993. .9655			
		/note="assembly_fragment"			
misc_feature		9756. .12205			
		/note="assembly_fragment"			
misc_feature		12306. .14310			
		/note="assembly_fragment"			
misc_feature		14411. .16707			
		/note="assembly_fragment"			
misc_feature		16808. .19279			
		/note="assembly_fragment"			
misc_feature		19380. .21937			
		/note="assembly_fragment"			
misc_feature		22038. .48726			
		/note="assembly_fragment"			
misc_feature		48827. .51847			
		/note="assembly_fragment"			
misc_feature		51948. .54626			
		/note="assembly_fragment"			
misc_feature		54727. .58944			
		/note="assembly_fragment"			
misc_feature		59045. .63316			
		/note="assembly_fragment"			
misc_feature		63417. .68257			
		/note="assembly_fragment"			
misc_feature		68358. .73280			
		/note="assembly_fragment"			
misc_feature		73381. .78209			
		/note="assembly_fragment"			
misc_feature		78310. .84542			
		/note="assembly_fragment"			
misc_feature		84643. .89899			
		/note="assembly_fragment"			
misc_feature		90000. .96909			
		/note="assembly_fragment"			
misc_feature		97010. .105708			
		/note="assembly_fragment"			
misc_feature		105809. .116467			
		/note="assembly_fragment"			
misc_feature		116568. .128407			
		/note="assembly_fragment"			
misc_feature		128508. .144528			
		/note="assembly_fragment"			
misc_feature		144629. .169246			
		/note="assembly_fragment"			
misc_feature		169347. .172241			
		/note="assembly_fragment"			
		clone_end:T7			
		vector_side:right"			
BASE COUNT	42562 a	41265 c	42571 g	43083 t	2760 others
ORIGIN					
Query Match					
		65.8%;		Score 20.4;	DB 2; Length 172241;

```
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ggtacctatcaggtagcgctgagcaatttg 31
   ||| ||||| ||| ||| |||||
Db 110832 GGTGCTATCAGCAAGCGATGACCAATAG 110803

RESULT 8
AC020552/c 178820 bp DNA linear PRI 09-MAY-2001
LOCUS AC020552 Homo sapiens BAC clone RP11-431K24 from 1, complete sequence.
DEFINITION AC020552
VERSION AC020552.4 GI:11067353
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 178820)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 178820)
AUTHORS Nguyen,C., Drone,K., Hawkins,M. and Podhrasky,A.
TITLE The sequence of Homo sapiens BAC clone RP11-431K24
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 178820)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 178820)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 178820)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 1, 2000 this sequence version replaced gi:7630992.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0431K24
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc
```


SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-12616. Actual start of this clone is at base position 1 of RP11-431K24; actual end is at base position 178820 of RP11-431K24.

The fidelity of the dinucleotide (GA) repeat from base position 124360 to 124387 cannot be guaranteed. The region at base position 134620 to 134624 is covered by a single subclone. The fidelity of sequence at base position 124152 to 124176 cannot be guaranteed. The regions from 123927 to 123930 and 145393 to 145403 are single-stranded and sequenced with only one chemistry.

FEATURES

source

1. .178820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-431K24"
/clone_lib="RPCI-11"

repeat_region

5. .71

/rpt_family="L1"

repeat_region

87. .364

/rpt_family="Alu"

repeat_region

1272. .1553

/rpt_family="Alu"

repeat_region

2779. .3084

/rpt_family="Alu"

repeat_region

3739. .4048

/rpt_family="Alu"

repeat_region

4229. .4444

/rpt_family="Alu"

repeat_region

4540. .4846

/rpt_family="MIR"

repeat_region

4886. .5014

/rpt_family="Alu"

repeat_region

5113. .5372

/rpt_family="ERV1"

misc_feature

6568. .7005

/note="similar to EST AI022484 (NID:g3237725) OW50C05.s1"

repeat_region

6720. .7031

/rpt_family="Alu"

repeat_region

7097. .7212

/rpt_family="L2"

misc_feature

7112. .7473

/note="similar to EST A1141387 (NID:g3648844) oy41f08.s1"

misc_feature

7158. .7444

/note="similar to EST AA843283 (NID:g2929801) aj16d01.s1"

repeat_region

7546. .7650

/rpt_family="L1"

repeat_region

7651. .8191

/rpt_family="ERV1"

repeat_region

8192. .8374

/rpt_family="L1"

repeat_region

8353. .8404

/rpt_family="L1"

repeat_region

8407. .8882

/rpt_family="ERV1"

repeat_region

8887. .9198

/rpt_family="Alu"

repeat_region

9202. .9464

/rpt_family="ERV1"

repeat_region

9471. .9595

repeat_region /rpt_family="L1"
9610. .9711
repeat_region /rpt_family="L1"
9720. .10077
repeat_region /rpt_family="L1"
10078. .10375
repeat_region /rpt_family="Alu"
10376. .10543
repeat_region /rpt_family="L1"
10997. .11038
repeat_region /rpt_family="MaLR"
11039. .11342
repeat_region /rpt_family="Alu"
11343. .11651
repeat_region /rpt_family="MaLR"
11775. .12137
repeat_region /rpt_family="MaLR"
12709. .13030
repeat_region /rpt_family="Alu"
13183. .13482
repeat_region /rpt_family="Alu"
13660. .13957
repeat_region /rpt_family="Alu"
14067. .14300
repeat_region /rpt_family="MaLR"
14301. .14602
repeat_region /rpt_family="Alu"
14603. .14782
repeat_region /rpt_family="MaLR"
14969. .15266
repeat_region /rpt_family="Alu"
15341. .15642
repeat_region /rpt_family="Alu"
15657. .15960
repeat_region /rpt_family="Alu"
16373. .16661
repeat_region /rpt_family="Alu"
16709. .16850
repeat_region /rpt_family="L1"
17101. .17178
repeat_region /rpt_family="ERV1"
17179. .17499
repeat_region /rpt_family="Alu"
17500. .17908
repeat_region /rpt_family="ERV1"
17921. .18009
repeat_region /rpt_family="Alu"
18064. .18195
repeat_region /rpt_family="Alu"
19470. .19669
repeat_region /rpt_family="L1"
19765. .20070
repeat_region /rpt_family="ERV1"
20369. .20670
repeat_region /rpt_family="Alu"
21550. .21850
repeat_region /rpt_family="Alu"
22158. .22462
repeat_region /rpt_family="Alu"
23533. .23834
repeat_region /rpt_family="Alu"
23886. .24012
repeat_region /rpt_family="Alu"
24308. .24637

Query Match 65.8%; Score 20.4; DB 9; Length 178820;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtagcgctgagcaatttg 31
||| ||||| ||| ||| |||||
Db 106290 GGTTCCTATCAGCAAGCGATGACCAATTAG 106261


```

----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162428 bases at least Q40
Consensus quality: 171987 bases at least Q30
Consensus quality: 176412 bases at least Q20
Insert size: 192000; agarose-fp
Insert size: 180254; sum-of-contigs
Quality coverage: 3.5 in Q20 bases; agarose-fp
Quality coverage: 3.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1085: contig of 1085 bp in length
* 1086 1185: gap of 100 bp
* 1186 2242: contig of 1057 bp in length
* 2243 2342: gap of 100 bp
* 2343 3408: contig of 1066 bp in length
* 3409 3508: gap of 100 bp
* 3509 3956: contig of 448 bp in length
* 3957 4056: gap of 100 bp
* 4057 5612: contig of 1556 bp in length
* 5613 5712: gap of 100 bp
* 5713 7899: contig of 2187 bp in length
* 7900 7999: gap of 100 bp
* 8000 9476: contig of 1477 bp in length
* 9477 9576: gap of 100 bp
* 9577 11577: contig of 2001 bp in length
* 11578 11677: gap of 100 bp
* 11678 13800: contig of 2123 bp in length
* 13801 13900: gap of 100 bp
* 13901 15510: contig of 1610 bp in length
* 15511 15610: gap of 100 bp
* 15611 17280: contig of 1670 bp in length
* 17281 17380: gap of 100 bp
* 17381 19596: contig of 2216 bp in length
* 19597 19696: gap of 100 bp
* 19697 21776: contig of 2080 bp in length
* 21777 21876: gap of 100 bp
* 21877 24032: contig of 2156 bp in length
* 24033 24132: gap of 100 bp
* 24133 26650: contig of 2518 bp in length
* 26651 26750: gap of 100 bp
* 26751 30914: contig of 4164 bp in length.
* 30915 31014: gap of 100 bp
* 31015 34039: contig of 3025 bp in length
* 34040 34139: gap of 100 bp
* 34140 37599: contig of 3460 bp in length
* 37600 37699: gap of 100 bp
* 37700 41746: contig of 4047 bp in length
* 41747 41846: gap of 100 bp
* 41847 46449: contig of 4603 bp in length
* 46450 46549: gap of 100 bp
* 46550 51055: contig of 4506 bp in length
* 51056 51155: gap of 100 bp
* 51156 55588: contig of 4433 bp in length
* 55589 55688: gap of 100 bp
* 55689 61100: contig of 5412 bp in length
* 61101 61200: gap of 100 bp
* 61201 65672: contig of 4472 bp in length
* 65673 65772: gap of 100 bp
* 65773 72053: contig of 6281 bp in length
* 72054 72153: gap of 100 bp
* 72154 76428: contig of 4275 bp in length
* 76429 76528: gap of 100 bp
* 76529 84178: contig of 7650 bp in length

```

misc_feature /note="assembly_fragment" 65773. .72053
misc_feature /note="assembly_fragment" 72154. .76428
/note="assembly_fragment" clone_end:SP6
vector_side:right"

Query Match 65.8%; Score 20.4; DB 2; Length 183754;
Best Local Similarity 80.0%; Pred. No. 78;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtagcgtgagcaatttg 31
|||||
Db 136461 GGTACCTATCTCGTAGAGCTGATAGATTG 136432

RESULT 11
AC024935/c
LOCUS
DEFINITION
AC024935
AC024935
VERSION
KEYWORDS
SOURCE
ORGANISM

AC024935 260430 bp DNA linear HTG 15-OCT-2001
Homo sapiens chromosome 12 clone RP11-471G7, WORKING DRAFT
SEQUENCE, 28 unordered pieces.
AC024935
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.

Homo saplens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 260430)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hagues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubenok,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
REFERENCE 2 (bases 1 to 260430)
Worley,K.C.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (03-MAR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 14, 2001 this sequence version replaced gi:11995488.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAJN
Center clone name: RP11-471G7
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-primer Bodipy: 1% of reads
Chemistry: Dye-terminator Big Dye: 99% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 260230 bases at least Q40
Consensus quality: 286983 bases at least Q30
Consensus quality: 302060 bases at least Q20
Estimated insert size: 285925; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 62329: contig of 62329 bp in length
* 62330 62429: gap of unknown length
* 62430 105693: contig of 43264 bp in length
* 105694 105793: gap of unknown length
* 105794 139148: contig of 33355 bp in length
* 139149 139248: gap of unknown length
* 139249 165659: contig of 26411 bp in length
* 165660 165759: gap of unknown length
* 165760 174283: contig of 8524 bp in length
* 174284 174383: gap of unknown length
* 174384 181214: contig of 6831 bp in length
* 181215 181314: gap of unknown length
* 181315 188422: contig of 7108 bp in length
* 188423 188522: gap of unknown length
* 188523 193685: contig of 5163 bp in length
* 193686 193785: gap of unknown length
* 193786 199658: contig of 5873 bp in length
* 199659 199758: gap of unknown length
* 199759 204510: contig of 4752 bp in length
* 204511 204610: gap of unknown length
* 204611 208759: contig of 4149 bp in length
* 208760 208859: gap of unknown length
* 208860 213510: contig of 4651 bp in length
* 213511 213610: gap of unknown length
* 213611 217169: contig of 3559 bp in length
* 217170 217269: gap of unknown length
* 217270 220435: contig of 3166 bp in length
* 220436 220535: gap of unknown length
* 220536 224670: contig of 4135 bp in length
* 224671 224770: gap of unknown length
* 224771 227815: contig of 3045 bp in length
* 227816 227915: gap of unknown length
* 227916 230995: contig of 3080 bp in length
* 230996 231096: gap of unknown length
* 231096 233174: contig of 2079 bp in length
* 233175 233274: gap of unknown length
* 233275 236639: contig of 3365 bp in length

REFERENCE 1 (sites)
AUTHORS Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and Nakase,T.
TITLE Phylogenetic analysis of saccharolytic oral treponemes isolated from human subgingival plaque
JOURNAL Microbiol. Immunol. 43 (7), 711-716 (1999)
MEDLINE 99456392
REFERENCE 2 (bases 1 to 1410)
AUTHORS Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and Nakase,T.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1998) to the DDBJ/EMBL/GenBank databases. Mitsuo Sakamoto, The Institute of Physical and Chemical Research (RIKEN), Japan Collection of Microorganisms; Hirosawa 2-1, Wako, Saitama 351-0198, Japan (E-mail:sakamoto@ulmus.riken.go.jp, Tel:81-48-462-1111(ex.5136), Fax:81-48-462-4619)
COMMENT On Jan 12, 1999 this sequence version replaced gi:4115455.
FEATURES Location/Qualifiers
 source 1..1410
 /organism="Treponema socranskii subsp. socranskii"
 /isolate="1166"
 /sub_species="socranskii"
 /db_xref="taxon:69715"
 <1..>1410
 /product="16S rRNA"
BASE COUNT 354 a 335 c 485 g 236 t
ORIGIN

Query Match 64.5%; Score 20; DB 1; Length 1410;
Best Local Similarity 82.1%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cgggtacctatcaggtagcgctgagcaat 28
||||| ||| ||||| ||||| |||
Db 1066 CGGTTACTACAGCTAGCGCTGAGGACT 1093

RESULT 15
AB015889
LOCUS 1414 bp DNA linear BCT 31-JUL-1999
DEFINITION Treponema socranskii subsp. socranskii gene for 16S rRNA, partial sequence, isolate:128B.
ACCESSION AB015889
VERSION AB015889.1 GI:4150942
KEYWORDS 16S ribosomal RNA.
SOURCE Treponema socranskii subsp. socranskii (sub_species:socranskii, isolate:128B) DNA.
ORGANISM Treponema socranskii subsp. socranskii
Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
REFERENCE 1 (sites)
AUTHORS Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and Nakase,T.
TITLE Phylogenetic analysis of saccharolytic oral treponemes isolated from human subgingival plaque
JOURNAL Microbiol. Immunol. 43 (7), 711-716 (1999)
MEDLINE 99456392
REFERENCE 2 (bases 1 to 1414)
AUTHORS Sakamoto,M., Koseki,T., Umeda,M., Ishikawa,I., Benno,Y. and Nakase,T.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1998) to the DDBJ/EMBL/GenBank databases. Mitsuo Sakamoto, The Institute of Physical and Chemical Research (RIKEN), Japan Collection of Microorganisms; Hirosawa 2-1, Wako, Saitama 351-0198, Japan (E-mail:sakamoto@ulmus.riken.go.jp, Tel:81-48-462-1111(ex.5136), Fax:81-48-462-4619)
COMMENT On Jan 12, 1999 this sequence version replaced gi:4115456.
FEATURES Location/Qualifiers
 source 1..1414
 /organism="Treponema socranskii subsp. socranskii"
 /isolate="128B"
 /sub_species="socranskii"
 /db_xref="taxon:69715"

rRNA <1..>1414
/product="16S rRNA"
BASE COUNT 354 a 337 c 484 g 239 t
ORIGIN

Query Match 64.5%; Score 20; DB 1; Length 1414;
Best Local Similarity 82.1%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cgggtacctatcaggtagcgctgagcaat 28
||||| ||| ||||| ||||| |||
Db 1069 CGGTTACTAACAGCTAGCGCTGAGGACT 1096

Search completed: August 31, 2002, 18:04:30
Job time: 15723 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:17:15 ; Search time 824.08 Seconds
(without alignments)
64.586 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtagcctatcaggtagcgctgagcaatttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	31	AAS17546	Human acetylcholin
C 2	31	100.0	5767	AAS17547	Plasmid pTM034. S
C 3	31	100.0	14446	AAS17548	Plasmid pTM036. S
C 4	20.4	65.8	1725	AAS17549	Synthetic human ac
5	20	64.5	256	AAC29843	Human secreted pro
6	20	64.5	449	ABA45007	Human breast cell
7	20	64.5	449	ABA55480	Human foetal liver
8	20	64.5	449	ABA55543	Human foetal liver
9	20	64.5	449	ABA25195	Probe #3661 for ge

10	20	64.5	449	22	ABA25248	Probe #3714 for ge
11	20	64.5	449	22	AAK29178	Human bone marrow
12	20	64.5	449	22	AAK29237	Human bone marrow
13	20	64.5	449	22	AAI13769	Probe #3702 for ge
14	20	64.5	449	22	AAI13830	Probe #3763 for ge
15	20	64.5	449	22	AAI35130	Probe #3816 used t
16	20	64.5	449	22	AAI03648	Probe #3639 used t
C 17	20	64.5	1897	22	AAI59644	Human polynucleoti
18	19	61.3	5476	22	ABA18966	Human nervous syst
19	19	61.3	5476	22	ABA18967	Human nervous syst
20	19	61.3	5476	22	AAL36446	Human musculoskele
21	19	61.3	5476	22	AAL36447	Human musculoskele
22	19	61.3	5476	22	AAS32794	Human genomic DNA
23	19	61.3	5476	22	AAS32795	Human genomic DNA
24	19	61.3	5476	22	AAS26816	Human genomic DNA
25	19	61.3	5476	22	AAS26817	Human genomic DNA
C 26	18.8	60.6	5471	24	ABI99407	Mouse ischaemic co
27	18.8	60.6	16075	20	AAV99811	Gum gene cluster.
28	18.8	60.6	16079	8	AAV99811	Sequence of segmen
29	18.2	58.7	47	21	AA287020	RBP-7 biallelic ma
30	18.2	58.7	1366	21	AAAI5481	16S rRNA sequence
C 31	18.2	58.7	4875	23	ABL16962	Drosophila melanog
32	18	58.1	1104	23	ABL20007	Drosophila melanog
33	18	58.1	3278	23	ABL20006	Drosophila melanog
C 34	17.8	57.4	879	21	AAC45814	Arabidopsis thalia
C 35	17.8	57.4	1035	21	AAC41930	Arabidopsis thalia
C 36	17.8	57.4	3573	23	ABL02260	Drosophila melanog
C 37	17.8	57.4	3948	23	ABL03664	Drosophila melanog
C 38	17.8	57.4	6065	22	AAS32886	Drosophila melanog
C 39	17.8	57.4	12113	24	AAS17492	Human genomic DNA
40	17.6	56.8	669	21	AAF12483	Human acetylcholin
C 41	17.6	56.8	7461	22	AAS30639	Aspergillus oryzae
C 42	17.6	56.8	7461	22	AAS28701	DNA encoding novel
C 43	17.6	56.8	32174	22	ABA15665	Genomic sequence #
C 44	17.6	56.8	32174	22	ABA19477	Human nervous syst
45	17.6	56.8	32174	22	ABA20359	Human nervous syst

ALIGNMENTS

RESULT 1
AAS17546
ID AAS17546 standard; DNA; 31 BP.
XX
AC AAS17546;
XX
DT 25-FEB-2002 (first entry)
XX
DE Human acetylcholinesterase, PCR primer AChE-Kpn.
XX
KW Human; acetylcholinesterase; AChE; PCR primer; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 4; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents PCR primer AChE-kpn used to amplify
CC human AChE as described in the method of the invention.
XX
SQ Sequence 31 BP; 7 A; 7 C; 9 G; 8 T; 0 other;

Query Match 100.0%; Score 31; DB 24; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgggtacctatcaggtagcgctgagcaatttg 31
Db 1 cgggtacctatcaggtagcgctgagcaatttg 31

RESULT 2
AAS17547/c
ID AAS17547 standard; DNA; 5767 BP.
XX
AC AAS17547;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM034.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.

PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX
DR WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 28-31; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM034, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;

Query Match 100.0%; Score 31; DB 24; Length 5767;
Best Local Similarity 100.0%; Pred. No. 3.8e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgggtacctatcaggtagcgctgagcaatttg 31
Db 2571 CGGTACCTATCAGGTAGCGCTGAGCAATTG 2541

RESULT 3
AAS17548/c
ID AAS17548 standard; DNA; 14446 BP.
XX
AC AAS17548;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM036.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM036;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.

```
XX 16-MAR-2001; 2001WO-US08468.
XX
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigamine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pRM036, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;

Query Match 100.0%; Score 31; DB 24; Length 14446;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtaccctatcaggtagcgctgagcaatttg 31
   |||||
Db 1875 CGGTACCTATCAGGTAGCGCTGAGCAATTG 1845

RESULT 4
AAS17549/c
ID AAS17549 standard; DNA; 1725 BP.
XX
AC AAS17549;
XX
DT 25-FEB-2002 (first entry)
XX
DE Synthetic human acetylcholinesterase gene.
XX
KW Human; acetylcholinesterase; AChE; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigamine bromide; ds.
XX
```

```
OS Homo sapiens.
OS Synthetic.
XX
XX WO200171014-A2.
XX
XX 27-SEP-2001.
XX
XX 16-MAR-2001; 2001WO-US08468.
XX
XX 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX
XX WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 41-42; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigamine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents synthetic human AChE gene optimised for
CC expression in plants.
XX
SQ Sequence 1725 BP; 273 A; 553 C; 571 G; 328 T; 0 other;

Query Match 65.8%; Score 20.4; DB 24; Length 1725;
Best Local Similarity 95.5%; Pred. No. 4.4;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 tcaggtagcgctgagcaatttg 31
   |||||
Db 1725 TCAGGTGGCGCTGAGCAATTG 1704

RESULT 5
AAC29843
ID AAC29843 standard; cDNA; 256 BP.
XX
AC AAC29843;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 33918.
XX
```


KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
PN 06-SEP-2000.
XX 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
XX (GEST) GENSET.
PA Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI; 2000-500381/45.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 33918; 71pp + CD-ROM; English.
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 256 BP; 55 A; 52 C; 93 G; 56 T; 0 other;

Query Match 64.5%; Score 20; DB 21; Length 256;
Best Local Similarity 82.1%; Pred. No. 4.7;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 cggctacatcatcaggtagcgctgagcaat 28
||||| ||||| || ||||| |
Db 213 cggtcacagatcaggcaggcgctgagcatt 240

RESULT 6
ABA45007
ID ABA45007 standard; DNA; 449 BP.
XX
AC ABA45007;
XX 01-FEB-2002 (first entry)
DT
DE Human breast cell single exon nucleic acid probe #3702.
XX
KW Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX Homo sapiens.
OS
XX WO200157271-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00652.

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX Claim 1; SEQ ID NO 3702; 327pp + sequence listing; English.
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 tacctatcaggtagcgctgagcaatttg 31
| ||| ||||| ||||| ||||
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 7
ABA55480
ID ABA55480 standard; DNA; 449 BP.
XX
AC ABA55480;
XX 01-FEB-2002 (first entry)
DT
XX Human foetal liver single exon nucleic acid probe #3785.
DE
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KW Homo sapiens.
XX WO200157277-A2.
XX
PN 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00669.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.

CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 tacctatcaggtagcgctgagcaatttg 31
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 10
ABA25248
ID ABA25248 standard; DNA; 449 BP.
XX
AC ABA25248;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #3714 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US006566.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
DR
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 1; SEQ ID No 3714; 530pp; English.
XX
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 tacctatcaggtagcgctgagcaatttg 31
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 11
AAK29178
ID AAK29178 standard; DNA; 449 BP.
XX
AC AAK29178;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 3735.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US006568.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 3735; 658pp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 tacctatcaggtagcgctgagcaatttg 31
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 12


```
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID No 3763; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 tacctatcaggtagcgctgagcaatttg 31
| | | | | | | | | | | | | | | | | |
Db 196 ttctcccaggtagcactgagcagtttg 223

RESULT 15
AAI35130
ID AAI35130 standard; DNA; 449 BP.
XX
AC AAI35130;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #3816 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID No 3816; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 449 BP; 100 A; 105 C; 113 G; 131 T; 0 other;

Query Match 64.5%; Score 20; DB 22; Length 449;
Best Local Similarity 82.1%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 tacctatcaggtagcgctgagcaatttg 31
| | | | | | | | | | | | | | | | | |
Db 196 ttctcccaggtagcactgagcagtttg 223

Search completed: August 31, 2002, 18:17:17
Job time: 16439 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:24:33 ; Search time 6064.26 Seconds
(without alignments)
68.995 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cgggtacctatcaggtagcgctgagcaatttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	21.4	69.0	419	10 BF082150	BF082150 QV0-BN004
2	19.8	63.9	574	12 BH400848	BH400848 AG-ND-143
3	19.8	63.9	687	12 BH398899	BH398899 AG-ND-159
4	19.8	63.9	820	10 BF786352	BF786352 602112868
5	19.8	63.9	866	12 CNS01Y2G	AL172465 Tetraodon
6	19.4	62.6	430	10 BG984091	BG984091 PM3-CN015
7	19.4	62.6	664	9 AF201206	AF201206 AF201206
8	19.4	62.6	668	10 BE370901	BE370901 601219331
9	19.4	62.6	776	10 BG821572	BG821572 602725163
10	19.2	61.9	646	10 BE534304	BE534304 601235365
11	19	61.3	316	9 BB852642	BB852642 BB852642
12	19	61.3	361	9 AU005717	AU005717 AU005717
13	19	61.3	379	10 BG228080	BG228080 ux47d04.x
14	19	61.3	390	10 BE634195	BE634195 uv70b03.y
15	19	61.3	390	10 BE986391	BE986391 UI-M-CG0p
16	19	61.3	395	10 BE692000	BE692000 uv70b03.x
17	19	61.3	403	9 AA959441	AA959441 vw59g04.s

18	61.3	417	12 AQ789846	AQ789846 HS_3091_B
19	61.3	429	9 AI850006	AI850006 UI-M-BG0-
20	61.3	437	9 AA170254	AA170254 ms87f10.r
21	61.3	447	9 AI854861	AI854861 UI-M-BH0-
22	61.3	459	9 AI850010	AI850010 UI-M-BG0-
23	61.3	465	9 AU005700	AU005700 AU005700
24	61.3	488	10 BG277664	BG277664 ux47d04.y
25	61.3	504	9 AI646695	AI646695 ub64f12.x
26	61.3	517	10 BF455638	BF455638 UI-M-CG0p
27	61.3	538	9 AI840147	AI840147 UI-M-AO0-
28	61.3	538	10 BF461699	BF461699 UI-M-CG0p
29	61.3	542	9 AU004943	AU004943 AU004943
30	61.3	543	9 AU006277	AU006277 AU006277
31	61.3	566	9 AV398885	AV398885 AV398885
32	61.3	577	9 AV402228	AV402228 AV402228
33	61.3	584	10 BE988805	BE988805 UI-M-CG0p
34	61.3	584	10 BF464864	BF464864 UI-M-CG0p
35	61.3	585	10 BG228490	BG228490 ux44b05.x
36	61.3	585	10 BE623027	BE623027 up51b04.x
37	61.3	600	10 BG808869	BG808869 2123-34 M
38	61.3	604	9 AV405785	AV405785 AV405785
39	61.3	605	9 BB619072	BB619072 BB619072
40	61.3	625	10 BF465972	BF465972 UI-M-CG0p
41	61.3	638	9 AV402654	AV402654 AV402654
42	61.3	645	9 AU006433	AU006433 AU006433
43	61.3	645	9 AV405236	AV405236 AV405236
44	61.3	646	9 AU006241	AU006241 AU006241
45	61.3	646	9 BB641643	BB641643 BB641643

ALIGNMENTS

RESULT 1
BF082150
LOCUS BF082150 419 bp mRNA linear EST 18-OCT-2000
DEFINITION QV0-BN0042-150900-389-e07 BN0042 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF082150
VERSION BF082150.1 GI:10875980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 419)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-QV0-BN0042-150900-389-e07&t3=2000-09-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 419.
Location/Qualifiers
1. 419
/organism="Homo sapiens"
/db_xref="taxon:9606"

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions.*

BASE COUNT 112 a 102 c 90 g 126 t
ORIGIN

Query Match 62.6%; Score 19.4; DB 10; Length 430;
Best Local Similarity 79.3%; Pred. No. 4.5e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtagcgctgagcaattt 30
| ||| || |||| | ||||| |||||
Db 206 GATACATAGCAGGTGCTTTGAGCAATT 234

RESULT 7
AF201206 664 bp mRNA linear EST 29-SEP-2000
LOCUS AF201206 Botryllus schlosseri EST Botryllus schlosseri cDNA, mRNA
DEFINITION sequence.

ACCESSION AF201206
VERSION AF201206.1 GI:10419580
KEYWORDS Botryllus schlosseri.
SOURCE Botryllus schlosseri
ORGANISM Botryllus schlosseri
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
Stolidobranchia; Botryllidae; Botryllus.

REFERENCE 1 (bases 1 to 664)
AUTHORS Meyer,K.A., Leonard,P.M., Rosenberg,G.H., Miller,R.D. and Loker
,E.S.

TITLE Characterization of ESTs from the Colonial Tunicate, Botryllus
schlosseri
JOURNAL Unpublished (2000)
COMMENT Contact: Loker ES

Parasitology
Biology, University of New Mexico
Castetter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 5508
Fax: 505 277 0304
Email: esloker@unm.edu.

FEATURES
source
1..664
/organism="Botryllus schlosseri"
/db_xref="taxon:30301"
/clone_lib="Botryllus schlosseri EST"

BASE COUNT 198 a 164 c 150 g 152 t
ORIGIN

Query Match 62.6%; Score 19.4; DB 9; Length 664;
Best Local Similarity 79.3%; Pred. No. 4.9e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 gtacctatcaggtagcgctgagcaatttg 31
| ||| ||||| ||| ||||| ||
Db 205 GTATGGATCAGGTAGCGGTCAGCAATCTG 233

RESULT 8
BE370901/c 668 bp mRNA linear EST 21-JUL-2000
LOCUS BE370901 601219331f1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3588294 5',
DEFINITION mRNA sequence.

ACCESSION BE370901
VERSION BE370901.1 GI:9316264
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 668)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8752 row: n column: 07
High quality sequence stop: 38.
Location/Qualifiers
1..668
/organism="Mus musculus"
/strain="C57BL/6J (f1)"
/db_xref="taxon:10090"
/clone="IMAGE:3588294"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary."
Stem cell origin.
/lab_host="DH10B"
/note="Organ: lung; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 321 a 180 c 95 g 72 t
ORIGIN

Query Match 62.6%; Score 19.4; DB 10; Length 668;
Best Local Similarity 79.3%; Pred. No. 4.9e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggtagcgctgagcaattt 30
| |||| | | | | | ||||| |
Db 144 GGTACCAATTATGCAGTGTGAGCAATGT 116

RESULT 9
BG821572/c 776 bp mRNA linear EST 22-MAY-2001
LOCUS BG821572 602725163f1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4864718 5',
DEFINITION mRNA sequence.

ACCESSION BG821572
VERSION BG821572.1 GI:14169159
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 776)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1728 row: n column: 15
High quality sequence stop: 762.
Location/Qualifiers
1..776
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4864718"

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 cggtagctatcaggtacgctgagcaa 27
||||| ||||| ||||| ||||| |||||
Db 32 CGGACCTTCCAGCAGCGCTGGCAA 6

RESULT 12
AU005717/c
LOCUS
DEFINITION AU005717 Bombyx mori p50(Daizo) Bombyx mori cDNA clone wv40011,
mRNA sequence.
ACCESSION AU005717
VERSION AU005717.1 GI:4163088
KEYWORDS EST.
SOURCE domestic silkworm.
ORGANISM Bombyx mori

REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
TITLE ; Bombycoidea; Bombycidae; Bombyx.
JOURNAL 1 (bases 1 to 361)
COMMENT Establishment of cDNA database of Bombyx mori
Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
Unpublished (1999)
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmitta@nirs.go.jp
PROJECT = 'CREST project by JST'.
Location/Qualifiers
1. .361

FEATURES
source
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/clone="wv40011"
/clone_lib="Bombyx mori p50(Daizo)"
BASE COUNT 105 a 98 c 81 g 77 t

Query Match 61.3%; Score 19; DB 9; Length 361;
Best Local Similarity 81.5%; Pred. No. 6.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ggtacatcaggtacgctgagcaat 28
||||| ||||| ||||| ||||| |||||
Db 98 GGTCCGATCATAGCGCTGAGGAAT 72

RESULT 13
BG228080
LOCUS
DEFINITION BG228080 379 bp mRNA linear EST 08-FEB-2001
ux47d04.x1 Soares_NMAX_maxillary_process Mus musculus cDNA clone
IMAGE:3513271 3' similar to SW:RB11_HUMAN P29374 RETINOBLASTOMA
BINDING PROTEIN 1 ;, mRNA sequence.
ACCESSION BG228080
VERSION BG228080.1 GI:12715595
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 379)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1393847

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 1.

FEATURES
source

Location/Qualifiers
1. .379
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3513271"
/clone_lib="Soares_NMAX_maxillary_process"
/tissue_type="maxillary process"
/lab_host="DH10B (phage-resistant)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 116 a 67 c 98 g 98 t

Query Match 61.3%; Score 19; DB 10; Length 379;
Best Local Similarity 81.5%; Pred. No. 6.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ggtacatcaggtacgctgagcaat 28
||||| ||||| ||||| ||||| |||||
Db 203 GGTGCTATCAGGAAGCGGTTATCAAT 229

RESULT 14
BE634195/c
LOCUS
DEFINITION BE634195 390 bp mRNA linear EST 25-AUG-2000
uv70b03.y1 Soares mouse 3NbMS Mus musculus cDNA clone IMAGE:3412493
5' similar to gb:S66427 RETINOBLASTOMA BINDING PROTEIN 1 (HUMAN);,
mRNA sequence.
ACCESSION BE634195
VERSION BE634195.1 GI:9916883
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 390)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1088305

Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .390
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3412493"
/clone_lib="Soares mouse 3NbMS"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 390)
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1393847

(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 93 a 102 c 73 g 122 t
ORIGIN

Query Match 61.3%; Score 19; DB 10; Length 390;
Best Local Similarity 81.5%; Pred. No. 6.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 2 ggtacctatcaggtagcgctgagcaat 28
||| ||||| ||||| ||||| |||||
Db 155 GGTGCCTATCAGGAAGCGGTTATCAAT 129

RESULT 15
BE986391
LOCUS BE986391 390 bp mRNA linear EST 05-OCT-2000
DEFINITION UI-M-CG0p-bgr-h-10-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
UI-M-CG0p-bgr-h-10-0-UI 3', mRNA sequence.
ACCESSION BE986391
VERSION BE986391.1 GI:10660675
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 390)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA-No.

FEATURES Location/Qualifiers
1..390
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-CG0p-bgr-h-10-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu.
TAG_SEQ=None found"

BASE COUNT 115 a 74 c 109 g 91 t 1 others
ORIGIN

Query Match 61.3%; Score 19; DB 10; Length 390;

Best Local Similarity 81.5%; Pred. No. 6.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 2 ggtacctatcaggtagcgctgagcaat 28
||| ||||| ||||| ||||| |||||
Db 282 GGTGCCTATCAGGAAGCGGTTATCAAT 308

Search completed: August 31, 2002, 15:24:39
Job time: 6382 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:27:36 ; Search time 153.1 Seconds
(without alignments)
49.736 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cgggtacctatcaggtagcgtgagcaatttg 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.8	60.6	16075	3	US-09-096-942-1
2	18.8	60.6	16075	3	US-09-096-867-1
3	18.2	58.7	829	4	US-08-998-416-495
4	18.2	58.7	1366	3	US-09-191-099-4
5	18	58.1	3016	2	US-08-318-826A-7
6	18	58.1	3016	2	US-08-370-156-5
7	18	58.1	3016	3	US-08-814-095-5
8	17.8	57.4	3096	2	US-08-318-826A-6
9	17.8	57.4	3096	2	US-08-370-156-3
10	17.8	57.4	3096	3	US-08-814-095-3
11	17.8	57.4	35060	3	US-08-814-095-7
12	17.4	56.1	1845	1	US-07-732-962A-1
13	17.4	56.1	1845	5	PCT-US92-06106-1
14	17.4	56.1	2256	2	US-08-318-826A-5
15	17.4	56.1	2256	2	US-08-370-156-1
16	17.4	56.1	2256	3	US-08-814-095-1
17	17.4	56.1	3499	4	US-08-857-076-43
18	17.2	55.5	1462	3	US-09-191-099-3
19	17.2	55.5	2133	3	US-08-808-032-1
20	17	54.8	3955	1	US-08-229-515A-14
21	17	54.8	3955	1	US-08-645-865-14
22	17	54.8	9045	3	US-09-121-321-1
23	17	54.8	9045	4	US-08-933-803A-1
24	17	54.8	50937	4	US-09-428-517-1
25	16.8	54.2	1325	2	US-08-632-470-50
26	16.6	53.5	595	6	5242807-1
27	16.6	53.5	1413	3	US-09-191-099-1
					Sequence 1, Appli

28	16.6	53.5	1503	4	US-08-943-571-1	Sequence 1, Appli
c 29	16.4	52.9	9785	1	US-08-319-387-1	Sequence 1, Appli
c 30	16.2	52.3	861	2	US-08-743-637B-166	Sequence 166, App
c 31	16.2	52.3	861	3	US-08-526-840B-166	Sequence 166, App
c 32	16.2	52.3	1920	1	US-08-746-789A-1	Sequence 1, Appli
c 33	16.2	52.3	3000	2	US-08-928-692-9	Sequence 9, Appli
c 34	16.2	52.3	3060	2	US-08-560-398-1	Sequence 1, Appli
c 35	16	51.6	50	4	US-09-390-867A-28	Sequence 28, Appli
c 36	16	51.6	50	4	US-09-548-260-28	Sequence 28, Appli
c 37	16	51.6	1038	4	US-09-004-838-128	Sequence 128, App
c 38	16	51.6	3494	4	US-09-139-802-200	Sequence 200, App
c 39	16	51.6	3645	2	US-08-663-112-1	Sequence 1, Appli
c 40	16	51.6	5427	3	US-09-009-913-2	Sequence 2, Appli
c 41	16	51.6	5510	3	US-09-009-913-3	Sequence 3, Appli
c 42	16	51.6	5667	3	US-09-009-913-4	Sequence 4, Appli
c 43	15.8	51.0	29	1	US-08-222-177A-460	Sequence 460, App
c 44	15.8	51.0	1224	2	US-08-210-762E-19	Sequence 19, Appli
c 45	15.8	51.0	1224	4	US-09-106-075A-18	Sequence 18, Appli

ALIGNMENTS

RESULT 1
US-09-096-942-1
; Sequence 1, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,942
; CURRENT FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16075
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; PUBLICATION INFORMATION:
US-09-096-942-1

Query Match 60.6%; Score 18.8; DB 3; Length 16075;
Best Local Similarity 90.9%; Pred. No. 5.7;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 10 tcaggtagcgtgagcaatttg 31
|||||
Db 11351 tcaggtagcgtgagcgatgtg 11372

RESULT 2
US-09-096-867-1
; Sequence 1, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,867


```

; OTHER INFORMATION: n = unknown
;
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1300)
; OTHER INFORMATION: n = unknown
;
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1313)
; OTHER INFORMATION: n = unknown
US-09-191-099-4

```

Query Match 58.7%; Score 18.2; DB 3; Length 1366;
Best Local Similarity 74.2%; Pred. NO. 6.5;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CGGTACCTATCAGGTAGCGCTGAGCAATTG 31
- - - - -
Db 1049 CAGTTGCTAACAGGTAAACCGCTGAGGACTCTG 1079

RESULT 5
US-08-318-826A-7/c
; Sequence 7, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharma
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; ;

US-08-318-826A-7

Query Match 58.1%; Score 18; DB 2; Length 3016;
Best Local Similarity 80.8%; Pred. No. 9.7;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 cctatcaggtagcgctgagcaatttg 31
 ||| | ||| ||| ||| |||
Db 1888 CCGAGCGGGTGGCGCTGAGCAATTG 1863

RESULT 6
 US-08-370-156-5/c
 ; Sequence 5, Application US/08370156
 ; Patent No. 5932780
 ; GENERAL INFORMATION:
 ; APPLICANT: Soreq, Hermona
 ; APPLICANT: Zakut, Haim
 ; APPLICANT: Shani, Moshe
 ; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
 ; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Reising, Ethington, Barnard & Perry
 ; STREET: P.O. Box 4390
 ; CITY: Troy
 ; STATE: Michigan
 ; COUNTRY: US
 ; ZIP: 48099
 ; COMPUTER READABLE FORM: .
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/370,156
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kohn, Kenneth I.
 ; REGISTRATION NUMBER: 30,955
 ; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (810) 689-3500
 ; TELEFAX: (810) 689-4071
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3016 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 160..2010
 ;
 US-08-370-156-5

```

Query Match          58.1%; Score 18; DB 2; Length 3016;
Best Local Similarity 80.8%; Pred. No. 9.7;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy 6 cctatcaggtagcgctgagcaatttg 31
 ||| ||| ||| ||| |||
Db 1888 CCGAGCGGTGGCGCTGAGCAATTG 1863

RESULT 7
US-08-814-095-5/c
; Sequence 5, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: zakut, Haim


```
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
; US-08-370-156-3

Query Match 57.4%; Score 17.8; DB 2; Length 3096;
Best Local Similarity 90.5%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 11 caggtagcgtgagcaatttg 31
| | | | | | | | | | | | | | |
Db 1883 CCGTGGCGCTGAGCAATTG 1863

RESULT 10
US-08-814-095-3/c
; Sequence 3, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced AChE
; DESCRIPTION: comprising exons 2, 3, 4 and 5 as well as the translated portion
; DESCRIPTION: of Intron 4 (readthrough)"
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: 160..1959
US-08-814-095-3

Query Match 57.4%; Score 17.8; DB 3; Length 3096;
Best Local Similarity 90.5%; Pred. No. 12;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 11 caggtagcgtgagcaatttg 31
| | | | | | | | | | | | | | |
Db 1883 CCGTGGCGCTGAGCAATTG 1863

RESULT 11
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Cosmid including ACHE
; DESCRIPTION: promotor, ACHE gene and ARS gene"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7q22
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
```

```
;
;
; NAME/KEY: exon
; LOCATION: 24090..25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "(translation start:
; OTHER INFORMATION: 24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 5
;
; FEATURE:
;
; NAME/KEY: terminator
; LOCATION: 27385..27387
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: 28008..28129
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 6
;
; FEATURE:
;
; NAME/KEY: terminator
; LOCATION: 28129..28131
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (34528..34895)
; OTHER INFORMATION: /function= "arsenite resistance
; OTHER INFORMATION: gene"
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 1
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (34092..34358)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 2
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (33779..33963)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 3
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (33493..33591)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 4
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (33297..33408)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 5
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (32959..33094)
```

```
;
;
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 6
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (32569..32628)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 7
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (32386..32468)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 8
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (31894..32080)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 9
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (31363..31534)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 10
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (31131..31284)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 11
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
;
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "ARS"
; OTHER INFORMATION: /number= 16
;
;
; US-08-814-095-7
;
; Query Match 57.4%; Score 17.8; DB 3; Length 35060;
; Best Local Similarity 90.5%; Pred. No. 21;
; Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; QY 11 caggtagcgctgagcaatttg 31
; ||| |||||
;
; Db 27174 CCGTGGCGCTGACCAATTG 27154
;
; RESULT 12
;
; US-07-732-962A-1/c
; ; Sequence 1, Application US/07732962A
; ; Patent No. 5248604
; ; GENERAL INFORMATION:
; ; APPLICANT: Fischer, Melr
; ; TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
; ; TITLE OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
; ; NUMBER OF SEQUENCES: 2
```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,962A
; FILING DATE: 19910722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
; US-07-732-962A-1

Query Match 56.1%; Score 17.4; DB 1; Length 1845;
Best Local Similarity 94.7%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtagcgctgagcaatttg 31
||| |||||
Db 1722 GGTGGCGCTGAGCAATTG 1704

RESULT 13
PCT-US92-06106-1/c
; Sequence 1, Application PC/TUS9206106
; GENERAL INFORMATION:
; APPLICANT: Fischer, Melr
; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06106
; FILING DATE: 19920722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
; PCT-US92-06106-1

Query Match 56.1%; Score 17.4; DB 5; Length 1845;
Best Local Similarity 94.7%; Pred. No. 17;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtagcgctgagcaatttg 31
||| |||||
Db 1722 GGTGGCGCTGAGCAATTG 1704

RESULT 14
US-08-318-826A-5/c
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens

; FEATURE:
; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4 and 6"
US-08-318-826A-5

Query Match 56.1%; Score 17.4; DB 2; Length 2256;
Best Local Similarity 94.7%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtagcgctgagcaatttg 31
||| |||||
Db 1881 GGTGGCGCTGAGCAATTG 1863

RESULT 15
US-08-370-156-1/c
; Sequence 1, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-156-1

Query Match 56.1%; Score 17.4; DB 2; Length 2256;
Best Local Similarity 94.7%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 ggtagcgctgagcaatttg 31
||| |||||
Db 1881 GGTGGCGCTGAGCAATTG 1863

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: August 31, 2002, 15:36:11 ; Search time 9324.54 Seconds
(without alignments)
65.083 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 qatatctgcagccatgqctagggcccccgc 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3595312

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: en_ba:*
16: en_fun:*
17: en_hum:*
18: en_in:*
19: en_mu:*
20: en_om:*
21: en_or:*
22: en_ov:*
23: en_pat:*
24: en_ph:*
25: en_pl:*
26: en_ro:*
27: en_sts:*
28: en_un:*
29: en_vi:*
30: en_htg_*
31: en_htg_*
32: en_htg_*
33: en_htg_*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100	100	1	...
2	95	95	95	2	...
3	90	90	90	3	...
4	85	85	85	4	...
5	80	80	80	5	...
6	75	75	75	6	...
7	70	70	70	7	...
8	65	65	65	8	...
9	60	60	60	9	...
10	55	55	55	10	...
11	50	50	50	11	...
12	45	45	45	12	...
13	40	40	40	13	...
14	35	35	35	14	...
15	30	30	30	15	...
16	25	25	25	16	...
17	20	20	20	17	...
18	15	15	15	18	...
19	10	10	10	19	...
20	5	5	5	20	...

ALIGNMENTS

RESULT 1

AX275252

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHOR

TITLE

AX275252 29 bp DNA
Sequence 1 from Patent WO0171014.
AX275252
AX275252.1 GI:16547672

synthetic construct.
synthetic construct
artificial sequence.

1 (sites)

Mor, T., Soreq, H., Arntzen, C. and Mason, H.

Expression of recombinant human acetylcholinesterase in transgenic plants

Patent: WO 0171014-A 1 27-SEP-2001;

BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; MOR,

Tsafrir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,

Hugh S. (US)

Location/Qualifiers

1.29

/organism="synthetic construct"

```
/db_xref="taxon:32630"
```

/note="primer pAChE-Nco, derived from human AChE gene and

BASE COUNT 5 a 11 c 8 g 5 t modified to introduce an Nco I restriction site"
ORIGIN

Query Match 100.0%; Score 29; DB 6; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gatattgcagccatggctagggcccccgc 29
|||||
Db 1 GATATCTGCAGCCATGGCTAGGCCCCCGC 29

RESULT 2
AF432517 AF432517 474 bp DNA linear PLN 30-JAN-2002
LOCUS
DEFINITION Avena sativa clone As3431 transposon mariner-like transposase
pseudogene, partial sequence.
ACCESSION AF432517
VERSION AF432517.1 GI:18419466
KEYWORDS
SOURCE oat.
ORGANISM Avena sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Poae; Avena.
REFERENCE 1 (bases 1 to 474)
AUTHORS Feschotte,C. and Wessler,S.R.
TITLE Mariner-like transposases are widespread and diverse in flowering
plants
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (1), 280-285 (2002)
MEDLINE 21642679
PUBMED 11756687
REFERENCE 2 (bases 1 to 474)
AUTHORS Feschotte,C. and Wessler,S.R.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2001) Botany, Unlversity of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source Location/Qualifiers
1. .474
/organism="Avena sativa"
/db_xref="taxon:4498"
/clone="As3431"
<1. .>474
/gene="transposase"
/pseudo
repeat_region <1. .>474
/transposon="mariner-like"
BASE COUNT 161 a 73 c 120 g 120 t
ORIGIN

Query Match 62.1%; Score 18; DB 8; Length 474;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 tgcagccatggctaggcc 24
|||||
Db 120 TGCAGCCATGGCTAGGCC 137

RESULT 3
AF432531 AF432531 485 bp DNA linear PLN 30-JAN-2002
LOCUS
DEFINITION Hordeum vulgare clone Hv2936 transposon mariner-like transposase
pseudogene, partial sequence.
ACCESSION AF432531
VERSION AF432531.1 GI:18419481
KEYWORDS
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Poae; Avena.
REFERENCE 1 (bases 1 to 474)
AUTHORS Feschotte,C. and Wessler,S.R.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2001) Botany, Unlversity of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source Location/Qualifiers
1. .474
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="Hv2936"
<1. .>485
/transposon="mariner-like"
BASE COUNT 166 a 76 c 116 g 127 t
ORIGIN

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 485)
AUTHORS Feschotte,C. and Wessler,S.R.
TITLE Mariner-like transposases are widespread and diverse in flowering
plants
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (1), 280-285 (2002)
MEDLINE 21642679
PUBMED 11756687
REFERENCE 2 (bases 1 to 485)
AUTHORS Feschotte,C. and Wessler,S.R.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2001) Botany, University of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source Location/Qualifiers
1. .485
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="Hv2936"
<1. .>485
/transposon="mariner-like"
repeat_region <1. .>485
/transposon="mariner-like"
gene <1. .>485
/gene="transposase"
/pseudo
BASE COUNT 168 a 73 c 115 g 129 t
ORIGIN

Query Match 62.1%; Score 18; DB 8; Length 485;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 tgcagccatggctaggcc 24
|||||
Db 120 TGCAGCCATGGCTAGGCC 137

RESULT 4
AF432532 AF432532 485 bp DNA linear PLN 30-JAN-2002
LOCUS
DEFINITION Hordeum vulgare clone Hv2939 transposon mariner-like transposase
pseudogene, partial sequence.
ACCESSION AF432532
VERSION AF432532.1 GI:18419482
KEYWORDS
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 485)
AUTHORS Feschotte,C. and Wessler,S.R.
TITLE Mariner-like transposases are widespread and diverse in flowering
plants
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (1), 280-285 (2002)
MEDLINE 21642679
PUBMED 11756687
REFERENCE 2 (bases 1 to 485)
AUTHORS Feschotte,C. and Wessler,S.R.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2001) Botany, University of Georgia, 2502 Plant
Sciences, Athens, GA 30602, USA
FEATURES
source Location/Qualifiers
1. .485
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="Hv2939"
<1. .>485
/transposon="mariner-like"
repeat_region <1. .>485
/transposon="mariner-like"
gene <1. .>485
/gene="transposase"
/pseudo
BASE COUNT 166 a 76 c 116 g 127 t
ORIGIN

ORIGIN

Query Match 62.1%; Score 18; DB 8; Length 485;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 tgcagccatggctaggcc 24
|||||

Db 120 TGCAGCCATGGCTAGGCC 137

RESULT 5

AC010395 AC010395 130615 bp DNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens chromosome 5 clone CTD-2089024, complete sequence.
AC010395
AC010395
AC010395.6 GI:10337627
HTG.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130615)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished

REFERENCE

2 (bases 1 to 130615)
DOE Joint Genome Institute.

AUTHORS

TITLE

JOURNAL

Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 130615)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (29-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT

On Sep 29, 2000 this sequence version replaced gi:7711435.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.7.

FEATURES

source

1. 130615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2089024"

BASE COUNT 39938 a 24537 c 25833 g 40307 t

ORIGIN

Query Match 62.1%; Score 18; DB 9; Length 130615;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggctagg 22
|||||

Db 50672 TCTGCAGCCATGGCTAGG 50689

RESULT 6

AC012610 AC012610 135248 bp DNA linear PRI 15-AUG-2001
DEFINITION Homo sapiens chromosome 5 clone CTC-540012, complete sequence.
AC012610
AC012610
AC012610.6 GI:15187228
HTG.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (31-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 135248)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (22-JAN-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 135248)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (15-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT

On Aug 15, 2001 this sequence version replaced gi:12331459.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of Sequence;
Estimated Total Number of Errors is 0.4.

FEATURES

source

1. 135248
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-540012"

BASE COUNT 41508 a 25271 c 26886 g 41583 t

ORIGIN

Query Match 62.1%; Score 18; DB 9; Length 135248;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggctagg 22
|||||

Db 55313 TCTGCAGCCATGGCTAGG 55330

RESULT 7

AX018091

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Sequence 48 from Patent WO9946374.
AX018091
AX018091.1 GI:10042542
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 906)
Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarsky,C.
Human nucleic acid sequences from prostate tumour tissue
Patent: WO 9946374-A 48 16-SEP-1999.
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)

FEATURES

source

1. 906
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 176 a 303 c 263 g 164 t

ORIGIN

Query Match 58.6%; Score 17; DB 6; Length 906;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||
Db 228 CTGCAGCCATGGCTAGG 244

RESULT 8
BC006151
LOCUS BC006151 1316 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, clone MGC:13170 IMAGE:3354369, mRNA, complete cds.
ACCESSION BC006151
VERSION BC006151.1 GI:13544031
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1316)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 18 Row: d Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source Location/Qualifiers
1..1316
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:13170 IMAGE:3354369"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
555..908
/codon_start=1
/product="Unknown (protein for MGC:13170)"
/protein_id="AAH06151.1"
/db_xref="GI:13544032"
/translation="MTVLEAVLEIQAITGSRLLSMVPGPARPPGSCWDPTQTRTWLL
SHTPRRRWISGLPRASCRLGEPPPLPYCDQAYGEELSIRHRETAWLSRIDTAWPGA
PGVKQARILGELLV"
BASE COUNT 269 a 419 c 387 g 241 t
ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 1316;

Best Local Similarity 100.0%; Pred. No. 28;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||
Db 85 CTGCAGCCATGGCTAGG 101

RESULT 9
AX014146
LOCUS AX014146 1467 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 11 from Patent WO9954447.
ACCESSION AX014146
VERSION AX014146.1 GI:10040593
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1467)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarsky,C.
TITLE Human nucleic acid sequences of bladder tumour tissue
JOURNAL Patent: WO 9954447-A 11 28-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source Location/Qualifiers
1..1467
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 307 a 463 c 432 g 265 t
ORIGIN

Query Match 58.6%; Score 17; DB 6; Length 1467;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||
Db 229 CTGCAGCCATGGCTAGG 245

RESULT 10
AX018186
LOCUS AX018186 1467 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 159 from Patent WO9946374.
ACCESSION AX018186
VERSION AX018186.1 GI:10042564
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1467)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarsky,C.
TITLE Human nucleic acid sequences from prostate tumour tissue
JOURNAL Patent: WO 9946374-A 159 16-SEP-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source Location/Qualifiers
1..1467
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 307 a 463 c 432 g 265 t
ORIGIN

Query Match 58.6%; Score 17; DB 6; Length 1467;
Best Local Similarity 100.0%; Pred. No. 27;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||

Db 229 CTGCAGCCATGGCTAGG 245

RESULT 11
AC010325/c
LOCUS
DEFINITION
Homo sapiens chromosome 19 clone CTD-2568A17, *** SEQUENCING IN
PROGRESS ***, 2 ordered pieces.
AC010325
AC010325.6 GI:15145658
VERSION
KEYWORDS
SOURCE
HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 166799)
AUTHORS
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE
Direct Submission
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 166799)
AUTHORS
DOE Joint Genome Institute.
TITLE
Direct Submission
JOURNAL
Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Aug 9, 2001 this sequence version replaced gi:13699497.

COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 52106: contig of 52106 bp in length
* 52107 52206: gap of unknown length
* 52207 166799: contig of 114593 bp in length.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 816110
Center clone name: CTD-2568A17

Summary Statistics
Consensus quality: 164399 bases at least Q40
Consensus quality: 165282 bases at least Q30
Consensus quality: 165704 bases at least Q20
Estimated insert size: 162230000; agarose-fp estimation
Estimated insert size: 165898; sum-of-contigs

estimation
estimation
estimation.
Location/Qualifiers
1. .166799
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2568A17"
BASE COUNT 39495 a 43287 c 44988 g 38927 t 102 others
ORIGIN

Query Match 58.6%; Score 17; DB 2; Length 166799;
Best Local Similarity 100.0%; Pred. No. 8.9;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||

Db 131945 CTGCAGCCATGGCTAGG 131929

RESULT 12
AC027602/c
LOCUS
DEFINITION
Homo sapiens chromosome 17 clone RP11-795B6 map 17, WORKING DRAFT
SEQUENCE, 12 unordered pieces.
AC027602
AC027602.4 GI:11178143
VERSION
KEYWORDS
SOURCE
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 217346)
AUTHORS
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE
Homo sapiens chromosome 17, clone RP11-795B6
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 217346)
AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collimore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczek, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
Direct Submission
JOURNAL
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT
On Nov 15, 2000 this sequence version replaced gi:11136831.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9166
Center clone name: 795_B_6
----- Summary Statistics
Sequencing vector: M13; M77815; 31% of reads
Sequencing vector: Plasmid; n/a; 69% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 210748 bases at least Q40
Consensus quality: 213655 bases at least Q30
Consensus quality: 215058 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 216246; sum-of-contigs
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 8149: contig of 8149 bp in length
8150 8249: gap of 100 bp
8250 9592: contig of 1343 bp in length
9593 9692: gap of 100 bp
9693 10733: contig of 1041 bp in length
10734 10833: gap of 100 bp
10834 13519: contig of 2686 bp in length
13520 13619: gap of 100 bp
13620 17510: contig of 3891 bp in length
17511 17610: gap of 100 bp
17611 24602: contig of 6992 bp in length
24603 24702: gap of 100 bp
24703 35434: contig of 10732 bp in length
35435 35534: gap of 100 bp
35535 124474: contig of 88940 bp in length
124475 124574: gap of 100 bp
124575 134664: contig of 10090 bp in length
134665 134764: gap of 100 bp
134765 162343: contig of 27579 bp in length
162344 162443: gap of 100 bp
162444 208917: contig of 46474 bp in length
208918 209017: gap of 100 bp
209018 217346: contig of 8329 bp in length.

FEATURES

source

1. .217346
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-795B6"
/clone_lib="RPC1-11 Human Male BAC"
1. .8149
/note="assembly_fragment"
clone_end:SP6
vector_side:left
8250. .9592
/note="assembly_fragment"
9693. .10733
/note="assembly_fragment"
10834. .13519
/note="assembly_fragment"
13620. .17510
/note="assembly_fragment"
17611. .24602
/note="assembly_fragment"
24703. .35434
/note="assembly_fragment"
35535. .124474
/note="assembly_fragment"
124575. .134664
/note="assembly_fragment"
134765. .162343
/note="assembly_fragment"
162444. .208917
/note="assembly_fragment"
209018. .217346
/note="assembly_fragment"
clone_end:T7
vector_side:right

BASE COUNT 55109 a 54154 c 53923 g 53053 t 1107 others
ORIGIN

Query Match 58.6%; Score 17; DB 2; Length 217346;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 6135 CTGCAGCCATGGCTAGG 6119

RESULT 13

AB026686

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AB026686
Physcomitrella patens mRNA for chlorophyll a/b-binding protein precursor, complete cds.

AB026686
chlorophyll a/b-binding protein precursor.

AB026686
Physcomitrella patens cDNA to mRNA.

AB026686
Physcomitrella patens

AB026686
Eukaryota; Viridiplantae; Streptophyta; Bryophyta;

AB026686
Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

AB026686
1 (sites)

AB026686
Kiyosue,T. and Wada,M.

AB026686
Three cDNA sequences for light inducible genes from Physcomitrella

AB026686
patens

AB026686
Unpublished (1999)

AB026686
2 (bases 1 to 964)

AB026686
Kiyosue,T.

AB026686
Direct Submission

AB026686
Submitted (24-APR-1999) Tomohiro Kiyosue, National Institute for

AB026686
Basic Biology (NIBB), Biological Regulation Division; Myoudaigai,

AB026686
Okazaki, Aichi 444-8585, Japan (E-mail:tkiyosue@nibb.ac.jp,

AB026686
Tel:81-564-55-7610, Fax:81-564-55-7611)

AB026686
Location/Qualifiers

AB026686
1. .964

AB026686
/organism="Physcomitrella patens"

AB026686
/db_xref="taxon:3218"

AB026686
40. .846

AB026686
/codon_start=1

AB026686
/product="chlorophyll a/b-binding protein precursor"

AB026686
/protein_id="BAA77273.1"

AB026686
/db_xref="GI:4757371"

AB026686
/translation="MATAATAMNSTVLAGOALLKPVSELSKVNAGEARVTMRKTVSK

AB026686
SSGSDSIWYGADRPKFLGPFSGTSPSYLNGEAGDYGWDYAGLSSDPETFARNRELEV

AB026686
IHARWAMLGALCLTPELLAKSGVKFGEAVWFKAGAQIFSEGGLDYLGNPSLVHAQSI

AB026686
LAIWACQVILMGAVEGYRVAGGFLGDVTDPIYGGSFDPGLGADDDPTFAELKVKKEIK

AB026686
NGRLAMFSMFGFFVQAIVTGKPLENLNDHLADPVANNAWAYATNFVPGN"

AB026686
BASE COUNT 206 a 280 c 262 g 216 t

AB026686
ORIGIN

AB026686
Query Match 55.2%; Score 16; DB 8; Length 964;

AB026686
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

AB026686
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AB026686
QY 3 tatctgcagccatggc 18

AB026686
|||||

AB026686
Db 29 TATCTGCAGCCATGGC 44

AB026686
RESULT 14

AB026686
CHKPRION

AB026686
LOCUS

AB026686
DEFINITION

AB026686
ACCESSION

AB026686
VERSION

AB026686
KEYWORDS

AB026686
SOURCE

AB026686
ORGANISM

AB026686
chicken.

AB026686
Gallus gallus

AB026686
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AB026686
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

AB026686
Phasianinae; Gallus.

AB026686
1 (bases 1 to 969)

AB026686
Gabriel,J.-M., Oesch,B., Kretzschmar,H., Scott,M. and Prusiner,S.B.

AB026686
Molecular cloning of a candidate chicken prion protein

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

AB026686
Proc. Natl. Acad. Sci. U.S.A. 89, 9097-9101 (1992)

```

MEDLINE 93028411 Location/Qualifiers
FEATURES
  source
    1. .969
    /organism="Gallus gallus"
    /strain="White Leghorn"
    /sub_species="domesticus"
    /db_xref="taxon:9031"
    /sex="male"
    /tissue_type="liver"
    /dev_stage="adult"
    /germline
    /tissue_lib="Clontech"
    <1. .34
    /gene="Prn-p"
    /number=1
    <1. .>969
    /gene="Prn-p"
    35. .>969
    /gene="Prn-p"
    /number=2
    <35. .36
    /gene="Prn-p"
    <35. .>969
    /gene="Prn-p"
    /product="prion protein"
    37. .858
    /gene="Prn-p"
    /codon_start=1
    /product="prion protein"
    /protein_id="AAC28970.1"
    /db_xref="GI:212611"
    /translation="MARLLTTCLLALLAACTDVALSKKGKPGSGGWGAGSHROP
SYPRPGYPHNPGYPHNPGYPHNPGYPHNPGYPHNPGYPHNPGYPHNPGYPGWGQGYN
PSSGGSYHNQPKPKPTNFKHVAGAAAGAVVGGGLGGYAMGRVMSGMNYHFDSPDEY
RWSSENSARYPNRVYRDYSSVPQDFVADCFNITVTEYSIGPAARKNTSEAVAAAN
QTEVEMENKVVTKVIREMCVQYREYRLASGIQLHPADTWLAVLLLLLTTLFAMH"
    37. .108
    /gene="Prn-p"
    109. .855
    /gene="Prn-p"
    /product="prion protein"
    859. .>969
    /gene="Prn-p"
    194 a 331 c 270 g 174 t
BASE COUNT 194 a 331 c 270 g 174 t
ORIGIN
    sig_peptide
    mat_peptide
    3'UTR
Query Match 55.2%; Score 16; DB 5; Length 969;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 gcagccatggctaggc 23
    |||||
Db 31 GCAGCCATGGCTAGGC 46
RESULT 15
AC064206/c
ID AC064206 standard; DNA; HTG; 1002 BP.
XX
AC AC064206;
XX
SV AC064206.1
XX
DT 26-APR-2000 (Rel. 63, Created)
DT 26-APR-2000 (Rel. 63, Last updated, version 1)
XX
DE Giardia intestinalis clone NJ3676 strain WB-C6, LOW-PASS SEQUENCE SAMPLING.
XX HTG; HTGS_PHASE0.
XX
OS Giardia intestinalis
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
```

```

XX [1]
RN 1-1002
RA Morrison H.G., McArthur A.G., Nixon J., Eakin N.Q., Kim U., Crocker M.C.,
RA Hinkle G., Holder M.E., Sogin M.L.;
RT "Giardia: a model for ancient eukaryotic genome analysis";
RL unpublished.
XX [2]
RN 1-1002
RA Nixon J., Morrison H.G., McArthur A.G., Eakin N.Q., Kim U., Crocker M.C.,
RA Hinkle G., Holder M.E., Sogin M.L.;
RT ;
RL Submitted (22-APR-2000) to the EMBL/GenBank/DBDJ databases.
RL Josephine Bay Paul Center for Comparative Molecular Biology and Evolution,
RL Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543-1015, USA
XX
CC * NOTE: This record contains 1 individual
CC * sequencing reads that have not been assembled into
CC * contigs. Runs of N are used to separate the reads
CC * and the order in which they appear is completely
CC * arbitrary. Low-pass sequence sampling is useful for
CC * identifying clones that may be gene-rich and allows
CC * overlap relationships among clones to be deduced.
CC * However, it should not be assumed that this clone
CC * will be sequenced to completion. In the event that
CC * the record is updated, the accession number will
CC * be preserved.
CC 1 1002: contig of 1002 bp in length.
XX
FH Key Location/Qualifiers
FH source
FT 1. .1002
FT /db_xref="taxon:5741"
FT /organism="Giardia intestinalis"
FT /strain="WB-C6"
FT /clone="NJ3676"
XX
SQ Sequence 1002 BP; 282 A; 204 C; 258 G; 257 T; 1 other;

Query Match 55.2%; Score 16; DB 33; Length 1002;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 atatctgcagccatgg 17
    |||||
Db 399 ATATCTGCAGCCATGG 384

Search completed: August 31, 2002, 22:34:43
Job time: 25112 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:18:19 ; Search time 824.51 Seconds
(without alignments)
60.388 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattctgcagccatgctaggccccgcg 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	29	AAS17545	Human acetylcholin
C 2	18	62.1	3023	AAS77493	DNA encoding novel
C 3	18	62.1	3684	AAS80007	DNA encoding novel
4	17	58.6	271	AAC00220	Human secreted pro
5	17	58.6	906	AAZ33470	Human prostate can
6	17	58.6	1452	AAI89867	Human polynucleoti
7	17	58.6	1466	AAF16096	Human prostate can
8	17	58.6	1467	AAZ24399	Human bladder tumo
9	16	55.2	473	AAA79626	Pinus radiata cell

C 10	16	55.2	4356	21	AAA79736	Pinus radiata cell
C 11	16	55.2	4481	10	AAQ90958	Sequence encoding
C 12	16	55.2	4487	13	AAQ20514	Encodes germ cell
C 13	15	51.7	51	22	AAL31467	Human SNP oligonuc
C 14	15	51.7	204	19	AAX10671	Human biallelic po
C 15	15	51.7	375	22	AAF66350	Novel human polynu
C 16	15	51.7	517	22	ABA61689	Human foetal liver
C 17	15	51.7	517	22	ABA29336	Probe #7802 for ge
C 18	15	51.7	517	22	AAK09991	Human brain expres
C 19	15	51.7	517	22	AAK35885	Human bone marrow
C 20	15	51.7	517	22	AAI17169	Probe #7102 for ge
C 21	15	51.7	517	22	AAI41601	Probe #10287 used
C 22	15	51.7	657	22	AAH52128	Human AFP protein
C 23	15	51.7	685	21	AAA63739	DNA encoding a pol
C 24	15	51.7	1215	21	AAC46923	Arabidopsis thalia
C 25	15	51.7	1360	24	ABI99523	Mouse ischaemic co
C 26	15	51.7	1921	23	AAS76387	DNA encoding novel
C 27	15	51.7	1994	21	AAA63740	DNA encoding a pol
C 28	15	51.7	2109	22	AAS02049	DNA encoding molec
C 29	15	51.7	2390	22	AAS45000	cdNA encoding nove
C 30	15	51.7	2439	23	ABL05705	Drosophila melanog
C 31	15	51.7	6454	22	AAI57932	Human polynucleoti
C 32	15	51.7	6742	22	AAI59718	Human polynucleoti
C 33	15	51.7	12143	23	ABL05704	Drosophila melanog
C 34	15	51.7	12807	22	AAI05769	Human reproductive
C 35	15	51.7	13165	22	ABA07140	Human pancreatic c
C 36	15	51.7	13165	22	AAK89348	Human digestive sy
C 37	15	51.7	16831	23	AAS59607	Propionibacterium
C 38	15	51.7	42488	22	AAK66772	Human immune/haema
C 39	14	48.3	22	13	AAQ27953	CA55 primer 2. Sy
C 40	14	48.3	24	13	AAQ35539	Spacer MRSYN20. S
C 41	14	48.3	24	13	AAQ35538	Spacer MRSYN19. S
C 42	14	48.3	36	21	AAA73841	Human bone marrow
C 43	14	48.3	107	18	AAT73876	Maize G9 core prom
C 44	14	48.3	107	20	AAX27552	G9 core promoter s
C 45	14	48.3	140	11	AAQ06660	Synthetic antifree

ALIGNMENTS

RESULT 1

AAS17545
ID AAS17545 standard; DNA; 29 BP.
XX
AC AAS17545;
XX
DT 25-FEB-2002 (first entry)
XX
DE Human acetylcholinesterase, PCR primer AChE-Nco.

XX
KW Human; acetylcholinesterase; AChE; PCR primer; antidote; pesticide;
transgenic plant; acetylcholinesterase poisoning; chemical warfare;
muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
organophosphate(OP)-modified AChE; pyridostigamine bromide; ss.

OS Homo sapiens.
OS Synthetic.

XX WO200171014-A2.

XX PD 27-SEP-2001.

XX PF 16-MAR-2001; 2001WO-US08468.

XX PR 17-MAR-2000; 2000US-190440P.

XX PA (MORT/) MOR T.

XX PA (SORE/) SOREQ H.

XX PA (ARNT/) ARNTZEN C.

XX PA (MASO/) MASON H.

XX PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 4; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents PCR primer AChE-Nco used to amplify
CC human AChE as described in the method of the invention.
XX
SQ Sequence 29 BP; 5 A; 11 C; 8 G; 5 T; 0 other;

Query Match 100.0%; Score 29; DB 24; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.6e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatattctgcagccatggtaggccccgc 29
|||||
Db 1 gatattctgcagccatggtaggccccgc 29

RESULT 2
AAS77493/c
ID AAS77493 standard; cDNA; 3023 BP.
XX
AC AAS77493;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #13297.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.
DR P-PSDB; ABG13306.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID NO 13297; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3023 BP; 1083 A; 698 C; 631 G; 611 T; 0 other;

Query Match 62.1%; Score 18; DB 23; Length 3023;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggctagg 22
|||||
Db 1362 TCTGCAGCCATGGCTAGG 1345

RESULT 3
AAS80007/c
ID AAS80007 standard; cDNA; 3684 BP.
XX
AC AAS80007;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #15811.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.
DR P-PSDB: ABG15820.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID NO 15811; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3684 BP; 1391 A; 856 C; 712 G; 725 T; 0 other;

Query Match 62.1%; Score 18; DB 23; Length 3684;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 tctgcagccatggctagg 22
|||||
Db 2628 TCTGCAGCCATGGCTAGG 2611

RESULT 4
AAC00220
ID AAC00220 standard; cDNA; 271 BP.
XX
AC AAC00220;
XX
DT 06-OCT-2000 (first entry)
DE Human secreted protein 5' EST, SEQ ID NO: 218.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG00214.

XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 218; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 271 BP; 53 A; 84 C; 76 G; 56 T; 2 other;

Query Match 58.6%; Score 17; DB 21; Length 271;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 ctgcagccatggctagg 22
|||||
Db 200 ctgcagccatggctagg 216

RESULT 5
AAZ33470
ID AAZ33470 standard; cDNA; 906 BP.
XX
AC AAZ33470;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human prostate cancer-associated EST 48.
XX
KW Expressed sequence tag; EST; prostate tumor; antitumor; treatment;
KW gene therapy; tissue specificity human; ss.
XX
OS Homo sapiens.
XX
PN DE19811193-A1.
XX
PD 16-SEP-1999.
XX
PF 10-MAR-1998; 98DE-1011193.
XX
PR 10-MAR-1998; 98DE-1011193.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX
DR WPI; 1999-519628/44.
DR P-PSDB; AAY48276, AAY48277, AAY48278, AAY48279, AAY48280, AAY48281,
DR AAY48282.
XX
PT New nucleic acid expressed at high level in prostatic tumor tissue and
PT encoded polypeptides, useful for treating cancer and screening for
PT therapeutic agents -
XX
PS Claim 1a; 107; 166pp; German.
XX
CC This invention describes novel nucleic acid sequences (A) that are
CC expressed at high level in prostatic tumor tissue and encode gene

CC products or their fragments. The products of the invention have
CC antitumor activity. Polypeptides (I) encoded by (A) are used: (i) for
CC identifying agents for treatment of prostatic cancer and (ii) for
CC therapy of prostate cancer, optionally where expressed by gene therapy
CC methods. (A) is also used to isolate full-length genes (for gene therapy)
CC and for recombinant production of (I), which can be used to raise
CC specific antibodies. (A) are identified by assembly of ESTs (expressed
CC sequence tags) before they are analyzed for expression pattern (tissue
CC specificity). This approach eliminates many of the false results, as
CC regards tissue specificity, associated with known methods that use
CC single (usually short) ESTs. AAZ33423-233476 represent expressed
CC sequence tags described in the method of the invention.
XX
SQ Sequence 906 BP; 176 A; 303 C; 263 G; 164 T; 0 other;

Query Match 58.6%; Score 17; DB 20; Length 906;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 228 ctgcagccatggctagg 244

RESULT 6
AAI89867
ID AAI89867 standard; cDNA; 1452 BP.
XX
AC AAI89867;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 9927.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO09936.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 1; SEQ ID NO 9927; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1452 BP; 291 A; 460 C; 429 G; 272 T; 0 other;

Query Match 58.6%; Score 17; DB 22; Length 1452;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 207 ctgcagccatggctagg 223

RESULT 7
AAF16096
ID AAF16096 standard; cDNA; 1466 BP.
XX
AC AAF16096;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:531.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587513/55.
DR P-PSDB; AAB56893.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 1; Page 993; 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

XX SQ Sequence 1466 BP; 318 A; 453 C; 426 G; 269 T; 0 other;

Query Match 58.6%; Score 17; DB 21; Length 1466;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 211 ctgcagccatggctagg 227
|||||

RESULT 8
AAZ24399
ID AAZ24399 standard; cDNA; 1467 BP.
XX
AC AAZ24399;
XX
DT 14-FEB-2000 (first entry)
XX
DE Human bladder tumour cDNA library derived EST 11.
XX
KW Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
KW treatment; gene therapy; EST; ss.
XX
OS Homo sapiens.
XX
PN DE19818619-A1.
XX
PD 28-OCT-1999.
XX
PF 21-APR-1998; 98DE-1018619.
XX
PR 21-APR-1998; 98DE-1018619.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-612028/53.
XX
PT New nucleic acid sequences expressed in bladder tumor tissue, and
PT derived polypeptides, for treatment of bladder tumor and identification
PT of therapeutic agents -
XX
PS Claim 3; Page 68-69; 132pp; German.
XX
CC This invention describes novel polypeptide fragments (I) and the
CC polynucleotides (II) that encode them that are highly expressed in a
CC human bladder tumour and which have cytostatic activity. (II) are used
CC for recombinant expression of (I) and to isolate complete genes. (I) are
CC used to identify agents suitable for treatment of bladder cancer, to
CC directly treat this form of cancer (including expression from gene
CC therapy vectors) or are used in a preparation for cancer treatment. (I)
CC is also used for the generation of specific antibodies. (II) are
CC identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, and
CC therefore reduces the number of failures associated with the fact that
CC ESTs from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AAZ43260-243309 represent expressed sequence tag (EST)
CC fragments isolated from a human bladder tumour cDNA library which encode
CC the proteins represented in AAY66143-Y66198.
XX
SQ Sequence 1467 BP; 307 A; 463 C; 432 G; 265 T; 0 other;

Query Match 58.6%; Score 17; DB 20; Length 1467;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 229 ctgcagccatggctagg 245
|||||

RESULT 9
AAA79626
ID AAA79626 standard; cDNA; 473 BP.
XX
AC AAA79626;
XX
DT 27-NOV-2000 (first entry)
XX
DE Pinus radiata cell signalling involved polynucleotide SEQ ID NO:427.
XX
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism; ss.
XX
OS Pinus radiata.
XX
PN WO200042171-A1.
XX
PD 20-JUL-2000.
XX
PF 11-JAN-2000; 2000WO-US00724.
XX
PR 12-JAN-1999; 99US-0228986.
PR 01-NOV-1999; 99US-0162866.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Strabala TJ, Nieuwenhuizen NJ;
XX WPI; 2000-476052/41.
XX
PT Isolated polynucleotide encoding a polypeptide involved in cell
PT signaling used for generating transgenic plants with modified responses
PT to external signals -
XX
PS Claim 1; Page 192; 527pp; English.
XX
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.
XX
SQ Sequence 473 BP; 105 A; 96 C; 118 G; 154 T; 0 other;

Query Match 55.2%; Score 16; DB 21; Length 473;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tatctgcagccatggc 18
|||||
Db 163 tatctgcagccatggc 178
|||||


```
RESULT 10
AAA79736
ID AAA79736 standard; cDNA; 4356 BP.
XX
AC AAA79736;
XX
DT 27-NOV-2000 (first entry)
XX
DE Pinus radiata cell signalling involved polynucleotide SEQ ID NO:888.
XX
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism; ss.
XX
OS Pinus radiata.
XX
PN WO200042171-A1.
XX
PD 20-JUL-2000.
XX
PF 11-JAN-2000; 2000WO-US00724.
XX
PR 12-JAN-1999; 99US-0228986.
PR 01-NOV-1999; 99US-0162866.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Strabala TJ, Nieuwenhuizen NJ;
XX
DR WPI; 2000-476052/41.
XX
PT Isolated polynucleotide encoding a polypeptide involved in cell
PT signaling used for generating transgenic plants with modified responses
PT to external signals.
XX
PS Claim 1; Page 440-442; 527pp; English.
XX
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.
XX
SQ Sequence 4356 BP; 1273 A; 824 C; 901 G; 1358 T; 0 other;

Query Match 55.2%; Score 16; DB 21; Length 4356;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tatctgcagccatggc 18
Db 135 tatctgcagccatggc 150
|||||

RESULT 11
AAN90958/c
ID AAN90958 standard; DNA; 4481 BP.
XX
AC AAN90958;
```

```
XX
DT 31-JAN-1990 (first entry)
XX
DE Sequence encoding germ cell alkaline phosphatase (ALP).
XX
KW Germ cell alkaline phosphatase gene; placental ALP; cancer treatment;
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT TATA_signal 368..374
FT exon /*tag= a
FT exon /*tag= b
FT exon 611..727
FT exon /*tag= c
FT exon 839..954
FT exon /*tag= d
FT exon 1147..1321
FT exon /*tag= e
FT exon 1399..1565
FT exon /*tag= f
FT exon 1813..1947
FT exon /*tag= g
FT exon 2047..2119
FT exon /*tag= h
FT exon 2248..2382
FT exon /*tag= i
FT exon 2464..2655
FT exon /*tag= j
FT exon 2875..2991
FT exon /*tag= k
FT exon 3114..3409
FT exon /*tag= l
XX
PN WO8908254-A.
XX
PD 08-SEP-1989.
XX
PF 08-FEB-1989; 89WO-US00567.
XX
PR 04-MAR-1988; 88US-0161138.
XX
PA (JOLL-) LA JOLLA CANCER RESEARCH FOUNDATION.
XX
PI Milan JL;
XX
DR WPI; 1989-278386/38.
DR P-PSDB; AAP91776.
XX
PT Germ cell alkaline phosphatase gene, protein and antibodies - used in
XX diagnosis and therapeutic treatment of certain cancers.
XX
PS Disclosure; fig 2; 31pp; English.
XX
CC Sequence contains 11 exons (tag b - ) which when spliced together code
CC for germ cell alkaline phosphatase (ALP). The DNA was obtd. from partial
CC MboI-digested human placental and spleenic DNA genomic libraries
CC constructed in Charon 28 lambda phage. The DNA can be used to make probes
CC for diagnosis or antibodies to treat germ cell tumours where the ALP is
CC expressed. Germ cell ALP is also termed the "Nagao isoenzyme" and is a
CC placental-ALP (PLAP)-like isoenzyme showing 98% homology with PLAP.
XX
SQ Sequence 4481 BP; 899 A; 1409 C; 1374 G; 799 T; 0 other;

Query Match 55.2%; Score 16; DB 10; Length 4481;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 gccatggctaggcccc 26
Db 1779 GCCATGGCTAGGCCCC 1764
|||||
```



```
RESULT 12
AAQ20514/c
ID AAQ20514 standard; DNA; 4487 BP.
XX
AC AAQ20514;
XX
DT 12-MAY-1992 (first entry)
XX
DE Encodes germ cell alkaline phosphatase.
XX
KW Placental ALP; cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT GC_signal 148..156
FT GC_signal /*tag= a
FT GC_signal 315..323
FT GC_signal /*tag= b
FT GC_signal /note= "exact match to SPL binding site consensus"
FT GC_signal 339..347
FT repeat_region /*tag= c
FT repeat_region 247..267
FT repeat_region /*tag= d
FT repeat_unit 247..251
FT repeat_unit /*tag= e
FT TATA_signal /rpt_type= DIRECT
FT TATA_signal 368..374
FT TATA_signal /note= "four copies in repeat region"
FT TATA_signal /*tag= f
FT exon 450..516
FT exon /*tag= g
FT intron /number= 1
FT intron 517..610
FT exon /*tag= h
FT exon 611..727
FT exon /*tag= i
FT intron /number= 2
FT intron 728..838
FT exon /*tag= j
FT exon 839..954
FT intron /*tag= k
FT intron /number= 3
FT intron 955..1146
FT exon /*tag= l
FT exon 1147..1321
FT intron /*tag= m
FT intron /number= 4
FT intron 1322..1398
FT exon /*tag= n
FT exon 1399..1571
FT intron /*tag= o
FT intron /number= 5
FT intron 1572..1818
FT exon /*tag= p
FT exon 1819..1953
FT intron /*tag= q
FT intron /number= 6
FT intron 1954..2052
FT exon /*tag= r
FT exon 2053..2125
FT intron /*tag= s
FT intron /number= 7
FT intron 2126..2253
FT exon /*tag= t
FT exon 2254..2388
FT intron /*tag= u
FT intron /number= 8
FT intron 2389..2469
FT exon /*tag= v
FT exon 2470..2661
```

```
FT /*tag= w
FT /number= 9
FT 2662..2880
FT /*tag= x
FT 2881..2997
FT /*tag= y
FT /number= 10
FT 2998..3119
FT /*tag= z
FT 3119..4487
FT /*tag= aa
FT /number= 11
FT 1024..1030
FT /*tag= ab
FT GC_signal 1615..1622
FT /*tag= ac
FT GC_signal 1731..1739
FT /*tag= ad
FT GC_signal 1757..1765
FT /*tag= ae
FT GC_signal 1989..1996
FT /*tag= af
FT GC_signal complement (2205..2211)
FT GC_signal /*tag= ag
FT GC_signal complement (2309..2316)
FT GC_signal /*tag= ah
FT GC_signal 3054..3062
FT /*tag= ai
FT GC_signal complement (3555..3561)
FT GC_signal /*tag= aj
FT GC_signal complement (3884..3890)
FT GC_signal /*tag= ak
FT GC_signal complement (4187..4193)
FT polyA_signal /*tag= al
FT 4228..4233
FT /*tag= am
XX
PN US5081227-A.
XX
PD 14-JAN-1992.
XX
PF 04-MAR-1988; 88US-0164138.
XX
PR 04-MAR-1988; 88US-0164138.
XX
PA (LJOL-) LA JOLLA CANCER RES.
XX
PI Millan JL;
XX
DR WPI; 1992-048357/06.
DR P-PSDB; AAR20527.
XX
PT New germ cell alkaline phosphatase and peptide fragments - giving
PT antibodies specific to germ cell alkaline phosphatase, used to
PT detect target, image germ cell and for site-directed therapy
XX
PS Disclosure; Fig 2; 10pp; English.
XX
CC Two different human genomic DNA libraries, one from placental tissue
CC and one from spleen tissue, were screened for the presence of
CC placental Alkaline Phosphatase (PALP)-related sequences. One clone
CC was present in both libraries and was chosen for further study. Its
CC sequence was compared to the known sequence of cDNA encoding PALP and
CC intron-exon boundaries identified by aligning the genomic sequence with
CC the cDNA. The new coding sequence is that of germ-line ALP.
CC The sequences listed in the Features Table as "GC_signals", although
CC highlighted in the specification, are not described in any detail
CC so there significance/function cannot be given.
XX
SQ Sequence 4487 BP; 900 A; 1411 C; 1377 G; 799 T; 0 other;
```

Query Match 55.2%; Score 16; DB 13; Length 4487;

Best Local Similarity 100.0%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 gccatggctaggcccc 26
| | | | | | | | | | | | | |
Db 1785 GCCATGGCTAGGCCCC 1770

RESULT 13
AAL31467/C
ID AAL31467 standard; DNA; 51 BP.
XX
AC AAL31467;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #4675.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX

OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX

PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -

Claim 1; Page 2731; 4143pp; English.

The present invention relates to oligonucleotides encoding polymorphic
variants of proteins related to amylases, amyloid proteins, angiotensin,
apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
histones, kinases, colony stimulating factors, complement related
proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
G-protein coupled receptors and thioesterases. The present sequence is
one such oligonucleotide. The oligonucleotides and the peptides encoded
by them may be used in the prevention, diagnosis and treatment of
diseases associated with inappropriate expression of the proteins listed
above. Disorders that may be prevented, diagnosed and/or treated include
multifactorial diseases with a genetic component, such as autoimmune
diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
systemic lupus erythematosus and Grave's disease), inflammation, cancer
(e.g. cancers of the bladder, brain, breast, colon and kidney,
leukaemia), diseases of the nervous system and an infection of pathogenic
organisms.

Sequence 51 BP; 14 A; 16 C; 12 G; 9 T; 0 other;

Query Match 51.7%; Score 15; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggct 19
| | | | | | | | | | | | | |
Db 47 TCTGCAGCCATGGCT 33

RESULT 14
AAX10671/C
ID AAX10671 standard; DNA; 204 BP.
XX
AC AAX10671;
XX
DT 30-MAR-1999 (first entry)
XX
DE Human biallelic polymorphic DNA fragment WI-6303.
XX

KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
XX

OS Homo sapiens.

XX WO9820165-A2.

PN 14-MAY-1998.

XX 05-NOV-1997; 97WO-US20313.

XX 06-NOV-1996; 96US-0030455.

PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Hudson T, Lander ES, Wang D;

PI WPI; 1998-286974/25.

PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease

XX Claim 1; Page 70; 310pp; English.

CC AAX10269-X12937 are human DNA fragments which contain biallelic
CC polymorphic markers which have been isolated using the primers
CC represented in AAX09121-X10268. The base occupying the polymorphic site
CC is indicated by the appropriate IUPAC-IUB ambiguity code. These fragments
CC can be used in methods for determining polymorphic forms in an individual
CC for use in e.g. forensics, paternity testing or for phenotypic typing for
CC diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC familial hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness; obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.

XX Sequence 204 BP; 47 A; 60 C; 43 G; 53 T; 1 other;

Query Match 51.7%; Score 15; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggct 19
| | | | | | | | | | | | | |
Db 118 TCTGCAGCCATGGCT 104

```
RESULT 15
AAF66350
ID AAF66350 standard; cDNA; 375 BP.
XX
AC AAF66350;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 2106.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 850; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 375 BP; 71 A; 140 C; 105 G; 58 T; 1 other;
```

Query Match

Best Local Similarity 51.7%; Score 15; DB 22; Length 375;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

4 atctgcagccatggc 18

|||||

Db

241 atctgcagccatggc 255

THIS PAGE BLANK (USPTO)

GenCore version 4.5
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:26:31 ; Search time 6064.54 Seconds
 (without alignments)
 64.541 Million cell updates/sec

Title: US-09-810-861B-1
 Perfect score: 29
 Sequence: 1 gatattgcagccatggctaggccccgc 29

Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0
 Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
 1: em_estba:*
 2: em_esthum:*
 3: em_estin:*
 4: em_estmu:*
 5: em_estov:*
 6: em_estpl:*
 7: em_estro:*
 8: em_htc:*
 9: gb_est1:*
 10: gb_est2:*
 11: gb_htc:*
 12: gb_gss:*
 13: em_gss_hum:*
 14: em_gss_inv:*
 15: em_gss_pln:*
 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	58.6	305	9	AW404201 UI-HF-BL0
2	17	58.6	311	9	AA380508 EST93479
3	17	58.6	323	10	BF803833 MR0-CI007
4	17	58.6	393	10	BE839073 RC0-FN014
5	17	58.6	422	9	AA401866 zv65c11.r
6	17	58.6	425	9	AW238889 xb29c10.y
7	17	58.6	440	10	BG140096 EST480538
8	17	58.6	440	10	BE251243 601116352
9	17	58.6	441	9	AA314757 EST186580
10	17	58.6	448	10	BF187245 EST443532
11	17	58.6	456	10	BE313471 601148429
12	17	58.6	496	10	BE254863 601111187
13	17	58.6	501	10	BM021984 ie74f02.y
14	17	58.6	508	10	BF796488 602259947
15	17	58.6	510	9	AW247247 2820405.5
16	17	58.6	530	10	BE252091 601113638
17	17	58.6	534	10	BF309847 601892311

c	18	58.6	539	12	BH190610	BH190610 TC3-49C21
	19	58.6	551	10	BM021473	BM021473 ie77g09.y
	20	58.6	553	10	BE279799	BE279799 601157110
	21	58.6	565	10	BE395544	BE395544 601309888
	22	58.6	569	9	AA990251	AA990251 ua62b08.r
	23	58.6	570	10	BM010014	BM010014 603630775
	24	58.6	574	9	AW248246	AW248246 2819868.5
	25	58.6	574	10	BE259913	BE259913 601153452
	26	58.6	577	10	BE256672	BE256672 601107754
	27	58.6	580	10	BE252950	BE252950 601117541
	28	58.6	600	10	BI908883	BI908883 603066474
	29	58.6	604	10	BE396866	BE396866 601289671
	30	58.6	605	10	BE883652	BE883652 601507332
	31	58.6	607	10	BG334821	BG334821 602460814
	32	58.6	610	10	BE266003	BE266003 601194716
	33	58.6	616	10	BE267403	BE267403 601189637
	34	58.6	617	10	BG890568	BG890568 EST516419
	35	58.6	618	10	BI179157	BI179157 EST520102
	36	58.6	618	10	BE280185	BE280185 601158395
	37	58.6	619	10	BE251761	BE251761 601112458
	38	58.6	625	10	BE261435	BE261435 601147194
	39	58.6	634	10	BF689428	BF689428 602186787
	40	58.6	635	10	BE382514	BE382514 601297929
	41	58.6	638	10	BF312637	BF312637 601898128
	42	58.6	639	10	BE264391	BE264391 601191478
	43	58.6	643	10	BE617089	BE617089 601441620
	44	58.6	644	10	BE900214	BE900214 601673151
	45	58.6	646	10	BE885936	BE885936 601507567

ALIGNMENTS

RESULT 1
 AW404201
 LOCUS UI-HF-BL0-abq-b-08-0-UI.r1 NIH_MGC_37 Homo sapiens cDNA clone
 DEFINITION IMAGE:3057351 5', mRNA sequence.
 ACCESSION AW404201
 VERSION AW404201.1 GI:6923258
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 305)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 cDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward.
 Location/Qualifiers
 1. 305
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3057351"
 /clone_lib="NIH_MGC_37"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /cell_line="MGC85"
 /lab_host="DH10B (LTI)"
 /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (1.5-2.5kb). Directionally cloned. Cells provided by Louis

M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 51 a 99 c 90 g 65 t
ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 305;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 105 CTGCAGCCATGGCTAGG 121

RESULT 2
AA380508
LOCUS
DEFINITION AA380508 311 bp mRNA linear EST 21-APR-1997
ACCESSION
VERSION AA380508.1 GI:2032826
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
1 (bases 1 to 311)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult

,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palanco,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
,Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
Venter,J.C.

TITLE
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence

JOURNAL
MEDLINE
COMMENT
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.

FEATURES
source
1. 311
/organism="Homo sapiens"
/db_xref="ATCC (inhost):184901"
/db_xref="taxon:9606"
/clone_lib="Supt cells"
/cell_type="T-lymphocyte"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 58 a 93 c 87 g 71 t 2 others
ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 311;

Best Local Similarity 100.0%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 175 CTGCAGCCATGGCTAGG 191

RESULT 3
BF803833/C
LOCUS
DEFINITION BF803833 323 bp mRNA linear EST 12-JAN-2001
ACCESSION
VERSION BF803833.1 GI:12132822
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
1 (bases 1 to 323)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
MEDLINE
COMMENT
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR0&t2=MR0-CI0075-
021100-003-a09&t3=2000-11-02&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 323.

FEATURES
source
1. 323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CI0075"
/dev_stage="Adult"
/note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 73 a 98 c 104 g 48 t
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 116 CTGCAGCCATGGCTAGG 100

RESULT 4
BE839073

LOCUS BE839073 393 bp mRNA linear EST 22-SEP-2000
DEFINITION RC0-FN0140-100700-022-a08 FN0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE839073
VERSION BE839073.1 GI:10271555
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 393)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC0-FN0140-100700-022-a08st3-2000-07-10st4-1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 393.
High quality sequence stop: 393.
FEATURES
source
1. .393
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0140"
/dev_stage="Adult"
/note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI ; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the PUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 76 a 129 c 108 g 80 t
ORIGIN
Query Match 58.6%; Score 17; DB 10; Length 393;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 ctgcagccatggctagg 22
|||||
Db 131 CTGCAGCCATGGCTAGG 147
RESULT 5
AA401866 422 bp mRNA linear EST 16-MAY-1997
LOCUS
DEFINITION zv65cl1.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:758516 5', mRNA sequence.
ACCESSION AA401866
VERSION AA401866.1 GI:2055885
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 422)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie ,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 407.
High quality sequence stop: 407.
FEATURES
Location/Qualifiers
source
1. .422
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:758516"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 80 a 129 c 123 g 90 t
ORIGIN
Query Match 58.6%; Score 17; DB 9; Length 422;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 ctgcagccatggctagg 22
|||||
Db 174 CTGCAGCCATGGCTAGG 190
RESULT 6
AW238889 425 bp mRNA linear EST 13-DEC-1999
LOCUS
DEFINITION xb29cl0.y1 NCI_CGAP_Lu31 Homo sapiens cDNA clone IMAGE:2577714 5', mRNA sequence.
ACCESSION AW238889
VERSION AW238889.1 GI:6571211
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 425)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: xb29cl0.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40RP from Gibco
High quality sequence stop: 425.
Location/Qualifiers

FEATURES
source

1. .425
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2577714"
/clone_lib="NCI_CGAP_Lu31"
/sex="male"
/dev_stage="fetal, 14 wk post-conception"
/lab_host="DH10B"
/note="Organ: lung, cell line; Vector: pCMV-SPORT6;
Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally, no
5' adaptor. Primer: Oligo dT. Full-length library
constructed by Life Technologies."

BASE COUNT 76 a 143 c 119 g 87 t
ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||

Db 134 CTGCAGCCATGGCTAGG 150

RESULT 7

BG140096

LOCUS

DEFINITION

EST480538 wild tomato pollen Lycopersicon pennellii cDNA clone
CLPP16M11 5' sequence, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Lycopersicon pennellii.
Lycopersicon pennellii.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 440)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
Generation of ESTs from wild tomato (L. pennellii) pollen
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES
source

1. .440
Location/Qualifiers
/organism="Lycopersicon pennellii"
/cultivar="TA56"
/db_xref="taxon:28526"
/clone="CLPP16M11"
/clone_lib="wild tomato pollen"
/tissue_type="pollen"
/dev_stage="pollen collected from open flowers"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Pollen was collected from open flowers from
L.pennellii TA56, and stored at -80 C until library
construction."

BASE COUNT 118 a 170 c 61 g 90 t 1 others
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 440;

Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 atctgcagccatggcta 20
|||||

Db 13 ATCTGCAGCCATGGCTA 29

RESULT 8

BE251243

LOCUS

DEFINITION

BE251243 601116352F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3356964 5',
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 440)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI63 row: g column: 13
High quality sequence stop: 440.

Location/Qualifiers

1. .440
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3356964"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 80 a 147 c 125 g 88 t
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 440;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||

Db 137 CTGCAGCCATGGCTAGG 153

RESULT 9

AA314757

LOCUS

DEFINITION

AA314757 EST186580 Colon carcinoma (HCC) cell line II Homo sapiens cDNA 5',
end, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI63 row: g column: 13
High quality sequence stop: 440.

Location/Qualifiers

1. .440
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3356964"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 80 a 147 c 125 g 88 t
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 440;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||

Db 137 CTGCAGCCATGGCTAGG 153

RESULT 9

AA314757

LOCUS

DEFINITION

AA314757 EST186580 Colon carcinoma (HCC) cell line II Homo sapiens cDNA 5',
end, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI63 row: g column: 13
High quality sequence stop: 440.

Location/Qualifiers

1. .440
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3356964"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 80 a 147 c 125 g 88 t
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 440;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 441)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other_ESTs: THC166913
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavetigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source Location/Qualifiers
1..441
/organism="Homo sapiens"
/db_xref="ATCC (inhost):111318"
/db_xref="taxon:9606"
/clone_lib="Colon carcinoma (HCC) cell line II"
/tissue_type="colon"
/cell_type="KM12C"
/cell_line="KM12C(HCC)-parental human colon carcinoma ;Dukes B2"
/note="Organ: colon; Vector: pBluescript SK-; Site_1: ECORI; Site_2: XhoI"
BASE COUNT 84 a 143 c 124 g 89 t 1 others
ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 441;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 ctgcagccatggctagg 22
|||||
Db 184 CTCGAGCCATGGCTAGG 200

RESULT 10
BF187245
LOCUS EST443532 potato stolon, Cornell University Solanum tuberosum cDNA
DEFINITION clone cSTA39J13 5' sequence, mRNA sequence.
ACCESSION BF187245
VERSION BF187245.1 GI:11069464
KEYWORDS EST.
SOURCE Solanum tuberosum
potato.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 448)
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R., Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan ,B., Bougri,O., Buell,C.R., Ronning,C.M., Tanksley,S.D. and Baker ,B.
TITLE Generation of ESTs from potato swelling stolons
JOURNAL Unpublished (1999)
COMMENT Contact: TIGR
The Institute for Genomic Research
Rockville, MD 20850, USA
Tel: (301)838-0200
Email: <http://www.tigr.org/tdb/potato/>
For clone information, please contact Research Genetics, Libraries Division, tel# 1-800-711-6195, email- cdna@resgen.com.

FEATURES
source Location/Qualifiers
1..448
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; RNA was supplied by Christian Bachem & Beatrix Horvath(Laboratory of Plant Breeding, Dept. of Plant Sciences, Wageningen University, The Netherlands). Total RNA was isolated from developing axillary buds of potato nodal stem cuttings cultured on medium for the introduction of tuber formation as described in Bachem et al. (Plant Journal 1996). Tissue samples were taken of stages corresponding to growing stolons and the early stages of tuber formation."
BASE COUNT 120 a 178 c 62 g 88 t
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 atctgcagccatggcta 20
|||||
Db 13 ATCTGCAGCCATGGCTA 29

RESULT 11
BE313471
LOCUS 601148429F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163854 5',
DEFINITION mRNA sequence.
ACCESSION BE313471
VERSION BE313471.1 GI:9133413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheraia; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 456)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

REFERENCE cDNA Library Preparation: Ling Hong/Rubin Laboratory
AUTHORS cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
TITLE DNA Sequencing by: Incyte Genomics, Inc.
JOURNAL Clone distribution: MGC clone distribution information can be
COMMENT found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: L1CM125 row: 1 column: 07
High quality sequence stop: 455.
Location/Qualifiers
1. .456
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3163854"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
82 a 154 c 128 g 92 t

BASE COUNT
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||

Db 144 CTGCAGCCATGGCTAGG 160

RESULT 12
BE254863
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BE254863 496 bp mRNA linear EST 13-JUL-2000
60111187F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3352124 5',
mRNA sequence.
BE254863
BE254863.1 GI:9125296
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 496)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM150 row: m column: 21
High quality sequence stop: 496.
Location/Qualifiers
1. .496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3352124"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 456;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||

Db 144 CTGCAGCCATGGCTAGG 160

RESULT 12
BE254863
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BE254863 496 bp mRNA linear EST 13-JUL-2000
60111187F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3352124 5',
mRNA sequence.
BE254863
BE254863.1 GI:9125296
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 496)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM150 row: m column: 21
High quality sequence stop: 496.
Location/Qualifiers
1. .496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3352124"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 496;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||

Db 105 CTGCAGCCATGGCTAGG 121

RESULT 13
BM021984
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM021984 501 bp mRNA linear EST 30-OCT-2001
ie74f02.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA 5', mRNA sequence.
BM021984
BM021984.1 GI:16536340
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 501)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
High quality sequence stop: 434.
Location/Qualifiers
1. .501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 496;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||

Db 105 CTGCAGCCATGGCTAGG 121

RESULT 13
BM021984
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM021984 501 bp mRNA linear EST 30-OCT-2001
ie74f02.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA 5', mRNA sequence.
BM021984
BM021984.1 GI:16536340
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 501)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
High quality sequence stop: 434.
Location/Qualifiers
1. .501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT
ORIGIN

BASE COUNT
ORIGIN

91 a 172 c 138 g 95 t

Query Match 58.6%; Score 17; DB 10; Length 496;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctgcagccatggctagg 22
|||||

Db 105 CTGCAGCCATGGCTAGG 121

RESULT 13
BM021984
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM021984 501 bp mRNA linear EST 30-OCT-2001
ie74f02.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA 5', mRNA sequence.
BM021984
BM021984.1 GI:16536340
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 501)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
High quality sequence stop: 434.
Location/Qualifiers
1. .501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dt priming. Size-selected by column
fractionlation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT
ORIGIN

90 a 174 c 139 g 98 t

Query Match 58.6%; Score 17; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 131 CTGCAGCCATGGCTAGG 147

RESULT 14
BF796488
LOCUS
DEFINITION
BF796488 508 bp mRNA linear EST 12-JAN-2001
602259947F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4343243 5',
mRNA sequence.
ACCESSION
BF796488
VERSION
BF796488.1 GI:12101542
KEYWORDS
EST.
SOURCE
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 508)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Location/Qualifiers
row: n column: 12
plate: LLAM9959
High quality sequence stop: 505.

FEATURES
source
1..508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4343243"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 92 a 172 c 143 g 101 t
ORIGIN

Query Match 58.6%; Score 17; DB 10; Length 508;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 160 CTGCAGCCATGGCTAGG 176

RESULT 15.
AW247247
LOCUS
DEFINITION
AW247247 510 bp mRNA linear EST 07-JAN-2000
2820405.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820405 5',
mRNA sequence.
ACCESSION
AW247247
VERSION
AW247247.1 GI:6590240
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 510)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Other_ESTs: 2820405.3prime
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LLNL at:
[www-bio.llnl.gov/dbbrp/image/image.html](http://www.bio.llnl.gov/dbbrp/image/image.html) Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross_match from University of Washington Genome Center
PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
<http://www.genome.washington.edu>
plate: LLCM4 row: B column: 22
High quality sequence stop: 458.

FEATURES
source
1..510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2820405"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University Of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 92 a 167 c 150 g 101 t
ORIGIN

Query Match 58.6%; Score 17; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggctagg 22
|||||
Db 180 CTGCAGCCATGGCTAGG 196

Search completed: August 31, 2002, 19:58:44
Job time: 16333 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 17:15:36 ; Search time 154.73 Seconds
(without alignments)
46.037 Million cell updates/sec

Title: US-09-810-861B-1
Perfect score: 29
Sequence: 1 gatattgcagccatggctagggccccgc 29

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14	48.3	22	1 US-08-104-073-13	Sequence 13, Appl
2	14	48.3	24	1 US-08-105-483-50	Sequence 50, Appl
C 3	14	48.3	24	1 US-08-105-483-51	Sequence 51, Appl
4	14	48.3	24	1 US-08-709-209-50	Sequence 50, Appl
C 5	14	48.3	24	1 US-08-709-209-51	Sequence 51, Appl
6	14	48.3	24	1 US-08-458-101-50	Sequence 50, Appl
C 7	14	48.3	24	1 US-08-458-101-51	Sequence 51, Appl
8	14	48.3	107	2 US-08-893-049-6	Sequence 6, Appl
9	14	48.3	292	1 US-08-230-929-2	Sequence 2, Appl
10	14	48.3	292	2 US-08-477-978-6	Sequence 6, Appl
C 11	14	48.3	1774	1 US-08-377-292-3	Sequence 3, Appl
12	14	48.3	2068	4 US-09-318-448-15	Sequence 15, Appl
13	14	48.3	2188	1 US-07-706-872-2	Sequence 2, Appl
C 14	14	48.3	2505	1 US-08-391-615-1	Sequence 1, Appl
15	13	44.8	18	4 US-08-718-388-18	Sequence 18, Appl
16	13	44.8	19	4 US-08-718-388-22	Sequence 22, Appl
C 17	13	44.8	28	3 US-09-009-156-12	Sequence 12, Appl
C 18	13	44.8	28	4 US-09-372-154-12	Sequence 12, Appl
19	13	44.8	32	3 US-08-722-719-25	Sequence 25, Appl
20	13	44.8	42	3 US-08-836-337-1	Sequence 1, Appl
C 21	13	44.8	194	2 US-08-634-797-48	Sequence 48, Appl
C 22	13	44.8	239	4 US-09-328-869-15	Sequence 15, Appl
C 23	13	44.8	239	4 US-09-629-774A-15	Sequence 15, Appl
C 24	13	44.8	244	4 US-09-034-205-29	Sequence 29, Appl
C 25	13	44.8	244	4 US-08-934-097A-29	Sequence 29, Appl
C 26	13	44.8	244	4 US-08-851-588-29	Sequence 29, Appl
C 27	13	44.8	244	4 US-09-677-218B-29	Sequence 29, Appl

C 28	13	44.8	244	4 US-09-677-192-29	Sequence 29, Appl
C 29	13	44.8	281	2 US-08-757-653-126	Sequence 126, App
30	13	44.8	281	2 US-08-757-653-132	Sequence 132, App
C 31	13	44.8	289	4 US-09-034-205-23	Sequence 23, Appl
C 32	13	44.8	289	4 US-08-934-097A-23	Sequence 23, Appl
C 33	13	44.8	289	4 US-08-851-588-23	Sequence 23, Appl
C 34	13	44.8	289	4 US-09-677-218B-23	Sequence 23, Appl
C 35	13	44.8	289	4 US-09-677-192-23	Sequence 23, Appl
36	13	44.8	691	2 US-08-365-486A-29	Sequence 29, Appl
37	13	44.8	691	4 US-08-880-342-29	Sequence 29, Appl
38	13	44.8	730	4 US-09-342-084-5	Sequence 5, Appl
C 39	13	44.8	803	1 US-08-157-235-1	Sequence 1, Appl
C 40	13	44.8	803	1 US-08-157-235-2	Sequence 2, Appl
C 41	13	44.8	803	1 US-08-157-235-3	Sequence 3, Appl
42	13	44.8	993	4 US-08-959-004-2	Sequence 2, Appl
43	13	44.8	1115	1 US-08-190-802A-19	Sequence 19, Appl
44	13	44.8	1115	4 US-08-477-346-19	Sequence 19, Appl
45	13	44.8	1115	4 US-08-473-089-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-08-104-073-13/c
; Sequence 13, Application US/08104073
; Patent No. 5589610
; GENERAL INFORMATION:
; APPLICANT: De Beuckeleer, Marc
; APPLICANT: Herdies, Lydia
; APPLICANT: Gossele, Veronique
; APPLICANT: Mariana, Celestina
; TITLE OF INVENTION: Stamen-Specific Promoters from Corn
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5589610west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104,073
; FILING DATE: 05-AUG-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP92/00275
; FILING DATE: 05-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91401787.6
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91400300.9
; FILING DATE: 07-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 8076.92USWO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-104-073-13

Query Match 48.3%; Score 14; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggct 19
|||||
Db 22 CTGCAGCCATGGCT 9

RESULT 2

US-08-105-483-50
; Sequence 50, Application US/08105483
; Patent No. 5494807

; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/105,483
; FILING DATE: 12-AUG-1993
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-105-483-50

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggc 18
|||||
Db 5 TCTGCAGCCATGGC 18

RESULT 3

US-08-105-483-51/c
; Sequence 51, Application US/08105483
; Patent No. 5494807

; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE

; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/105,483
; FILING DATE: 12-AUG-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-105-483-51

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 tctgcagccatggc 18
|||||
Db 24 TCTGCAGCCATGGC 11

RESULT 4

US-08-709-209-50
; Sequence 50, Application US/08709209
; Patent No. 5762938

; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,209
; FILING DATE: 21-AUG-1996
; CLASSIFICATION: 424

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-709-209-50

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctgcagccatggc 18
|||||
Db 5 TCTGCAGCCATGGC 18

RESULT 5
US-08-709-209-51/c
; Sequence 51, Application US/08709209
; Patent No. 5762938
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 462
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,209
; FILING DATE: 21-AUG-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

;
; TOPOLOGY: linear
; US-08-709-209-51

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctgcagccatggc 18
|||||
Db 24 TCTGCAGCCATGGC 11

RESULT 6
US-08-458-101-50
; Sequence 50, Application US/08458101
; Patent No. 5766599
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Perkus, Marion E.
; APPLICANT: Taylor, Jill
; APPLICANT: Tartaglia, James
; APPLICANT: No. 5766599ton, Elizabeth K.
; APPLICANT: Riviere, Michel
; APPLICANT: de Taisne, Charles
; APPLICANT: Limbach, Keith J.
; APPLICANT: Johnson, Gerard P.
; APPLICANT: Pincus, Steven E.
; APPLICANT: Cox, William I.
; APPLICANT: Audonnet, Jean-Christophe Francis
; APPLICANT: Gettig, Russell Robert
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
; TITLE OF INVENTION: STRAIN
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; ADDRESSEE: c/o William S. Frommer
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,101
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2740
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-458-101-50

Query Match 48.3%; Score 14; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctgcagccatggc 18
|||||

Db 5 TCTGCAGCCCATGGC 18

RESULT 7

US-08-458-101-51/C

; Sequence 51, Application US/08458101

; Patent No. 5766599

; GENERAL INFORMATION:

; APPLICANT: Paoletti, Enzo

; APPLICANT: Perkus, Marion E.

; APPLICANT: Taylor, Jill

; APPLICANT: Tartaglia, James

; APPLICANT: No. 5766599ton, Elizabeth K.

; APPLICANT: Riviere, Michel

; APPLICANT: de Taisne, Charles

; APPLICANT: Limbach, Keith J.

; APPLICANT: Johnson, Gerard P.

; APPLICANT: Pincus, Steven E.

; APPLICANT: Cox, William I.

; APPLICANT: Audonnet, Jean-Christophe Francis

; APPLICANT: Gettig, Russell Robert

; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE

; TITLE OF INVENTION: STRAIN

; NUMBER OF SEQUENCES: 467

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford

; ADDRESSEE: c/o William S. Frommer

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/458,101

; FILING DATE: 01-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Frommer, William S.

; REGISTRATION NUMBER: 25,506

; REFERENCE/DOCKET NUMBER: 454310-2740

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 51:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 24 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-458-101-51

Query Match 48.3%; Score 14; DB 1; Length 24;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tctgcagcccatggc 18

|||||

Db 24 TCTGCAGCCCATGGC 11

RESULT 8

US-08-893-049-6

; Sequence 6, Application US/08893049

; Patent No. 5962769

; GENERAL INFORMATION:

; APPLICANT: HOWARD, John A.

; APPLICANT: ALBERTSEN, Marc C.

; TITLE OF INVENTION: INDUCTION OF MALE STERILITY IN PLANTS BY

; TITLE OF INVENTION: EXPRESSION OF HIGH LEVELS OF AVIDIN

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/893,049

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/475,582

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 33229/314/PIHI

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 107 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-893-049-6

Query Match 48.3%; Score 14; DB 2; Length 107;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 tatctgcagcccatg 16

|||||

Db 94 TATCTGCAGCCCATG 107

RESULT 9

US-08-230-929-2

; Sequence 2, Application US/08230929

; Patent No. 5470359

; GENERAL INFORMATION:

; APPLICANT: Huffman, Gary A.

; TITLE OF INVENTION: REGULATORY ELEMENT CONFERRING TAPETUM

; TITLE OF INVENTION: SPECIFICITY

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/230,929

; FILING DATE: 21-APR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/210/PIHI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-230-929-2

Query Match 48.3%; Score 14; DB 1; Length 292;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tatctgcagccatg 16
|||||

Db 279 TATCTGCAGCCATG 292

RESULT 10
US-08-477-978-6
; Sequence 6, Application US/08477978
; Patent No. 5837850
; GENERAL INFORMATION:
; APPLICANT: HUFFMAN, Gary A.
; TITLE OF INVENTION: REGULATORY ELEMENT CONFERRING TAPETUM
; TITLE OF INVENTION: SPECIFICITY
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,978
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/230,929
; FILING DATE: 21-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/330/PIHI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-477-978-6

Query Match 48.3%; Score 14; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tatctgcagccatg 16
|||||

Db 279 TATCTGCAGCCATG 292

RESULT 11
US-08-377-292-3/c
; Sequence 3, Application US/08377292
; Patent No. 5693615
; GENERAL INFORMATION:
; APPLICANT: STONE, ROGER L.
; TITLE OF INVENTION: THERAPEUTIC FORMULAS FOR OSTEOINDUCTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: U.S.A.
; ZIP: 45239-8707
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,292
; FILING DATE: 23-JAN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,435
; FILING DATE:
; APPLICATION NUMBER: US/08/117,367
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corstanje, Brahm J.
; REGISTRATION NUMBER: 34,804
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-245-2858
; TELEFAX: 513-741-3012
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1774 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-377-292-3

Query Match 48.3%; Score 14; DB 1; Length 1774;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 cagccatggctagg 22
|||||

Db 327 CAGCCATGGCTAGG 314

RESULT 12
US-09-318-448-15
; Sequence 15, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stenroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2068
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-15

Query Match 48.3%; Score 14; DB 4; Length 2068;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggct 19
|||||
Db 68 ctgcagccatggct 81

RESULT 13

US-07-706-872-2
; Sequence 2, Application US/07706872
; Patent No. 5237056
; GENERAL INFORMATION:

; APPLICANT: Fischbach, Gerald D.
; TITLE OF INVENTION: Identification of a Protein Which
; TITLE OF INVENTION: Promotes the Synthesis of Acetylcholine Receptors and Uses
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07706,872
; FILING DATE: 19910529
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU91-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2188 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 172..972
US-07-706-872-2

Query Match 48.3%; Score 14; DB 1; Length 2188;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 agccatggctaggc 23
|||||
Db 168 AGCCATGGCTAGGC 181

RESULT 14

US-08-391-615-1/c
; Sequence 1, Application US/08391615
; Patent No. 5550054
; GENERAL INFORMATION:

; APPLICANT: Witte, Owen
; APPLICANT: Tsukada, Satoshi
; APPLICANT: Saffran, Douglas
; APPLICANT: Rawlings, David
; TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,615
; FILING DATE:
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/006,449
; FILING DATE: 21-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2505 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 137..2116
US-08-391-615-1

Query Match 48.3%; Score 14; DB 1; Length 2505;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctgcagccatggct 19
|||||
Db 146 CTGCAGCCATGGCT 133

RESULT 15

US-08-718-388-18
; Sequence 18, Application US/08718388
; Patent No. 6271362
; GENERAL INFORMATION:

; APPLICANT: MORIKAWA, MINORU
; APPLICANT: HARADA, NAOKI
; TITLE OF INVENTION: GENE ENCODING IgG Fc REGION-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747

;
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,388
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0230-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "oligonucleotide"
US-08-718-388-18

Query Match 44.8%; Score 13; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 tctgcagccatgg 17
|||
Db 5 TCTGCAGCCATGG 17

Search completed: August 31, 2002, 22:37:20
Job time: 19304 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: August 31, 2002, 22:34:43 ; Search time 9324.54 Seconds
(without alignments)
69.572 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtaacctatcaggtagcgtgagcaatttg 31

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description

1	31	100.0	31	6	AX275253	AX275253 Sequence
C 2	31	100.0	5767	6	AX275254	AX275254 Sequence
C 3	31	100.0	14446	6	AX275255	AX275255 Sequence
4	18	58.1	23223	2	AC005975	AC005975 Drosophil
5	18	58.1	146775	2	AC013153	AC013153 Drosophil
6	18	58.1	175451	3	AC099035	AC099035 Drosophil
C 7	18	58.1	181463	3	AC008352	AC008352 Drosophil
C 8	18	58.1	299751	3	AE003459	AE003459 Drosophil
9	17	54.8	34690	3	AF099920	AF099920 Caenorhab
10	17	54.8	178881	2	AC044838	AC044838 Mus muscu
11	17	54.8	205165	2	AL606531	AL606531 Mus muscu
12	17	54.8	209365	2	AC006719	AC006719 Caenorhab
C 13	17	54.8	271254	2	AL645602	AL645602 Mus muscu
14	16	51.6	1019	1	DSP276564	AJ276564 Desulfoto
C 15	16	51.6	1585	9	HSU34962	U34962 Human trans
16	16	51.6	2235	8	AB027318	AB027318 Cordyceps
17	16	51.6	2958	8	AB027322	AB027322 Cordyceps
18	16	51.6	16075	1	XCU22511	U22511 Xanthomonas
19	16	51.6	89172	8	AC084295	AC084295 Oryza sat
20	16	51.6	89813	2	AC023660	AC023660 Homo sapi
C 21	16	51.6	113392	2	AC099323	AC099323 Oryza sat
C 22	16	51.6	130049	8	NC17E5	AL513467 Neurospor
23	16	51.6	136657	2	AL669924	AL669924 Homo sapi
C 24	16	51.6	137545	2	AC008412	AC008412 Homo sapi
C 25	16	51.6	163275	2	AC007623	AC007623 Homo sapi
26	16	51.6	170518	2	AC091787	AC091787 Oryza sat
27	16	51.6	213891	2	AC007834	AC007834 Homo sapi
C 28	16	51.6	225506	2	AC102953	AC102953 Homo sapi
C 29	15	48.4	510	9	HUMACHE03	L22561 Human acety
C 30	15	48.4	535	9	HUMACHEA	M76539 Human acety
C 31	15	48.4	840	10	MUSACHEA	M76540 Mouse acety
C 32	15	48.4	1024	8	THQID74PR	Y16305 Trichoderma
C 33	15	48.4	1024	8	TRQID74PR	Y16306 Trichoderma
C 34	15	48.4	1415	9	HSU92014	U92014 Human clone
C 35	15	48.4	1641	9	AF278605	AF278605 Homo sapi
C 36	15	48.4	1725	6	AX275256	AX275256 Sequence
C 37	15	48.4	1838	8	AF007533	AF007533 Coemansia
38	15	48.4	2050	8	AB027512	AB027512 Lentinula
39	15	48.4	2051	6	E16408	E16408 Lentinus ed
C 40	15	48.4	2066	10	S50879	S50879 acetylcholi
C 41	15	48.4	2089	6	AX306139	AX306139 Sequence
C 42	15	48.4	2089	10	MMACHE	X56518 Mouse mRNA
C 43	15	48.4	2218	9	HUMACHE	M55040 Human acety
C 44	15	48.4	2256	6	AR070205	AR070205 Sequence
C 45	15	48.4	2281	9	AC002251	AC002251 Homo sapi

ALIGNMENTS

RESULT 1	AX275253	Sequence 2 from Patent WO0171014.	31 bp	DNA	linear	PAT 29-OCT-2001
AX275253	Sequence 2 from Patent WO0171014.					
LOCUS	AX275253					
DEFINITION	AX275253					
ACCESSION	AX275253					
VERSION	AX275253.1	GI:16547673				
KEYWORDS						
SOURCE		synthetic construct.				
ORGANISM		synthetic construct				
		artificial sequence.				
REFERENCE		1 (sites)				
AUTHORS		Mor,T., Soreq,H., Arntzen,C. and Mason,H.				
TITLE		Expression of recombinant human acetylcholinesterase in transgenic plants				
JOURNAL		Patent: WO 0171014-A 2 27-SEP-2001; BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor, Tsafirir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason, Hugh S. (US)				
FEATURES		Location/Qualifiers				
source		1. .31				
		/organism="synthetic construct"				
		/db_xref="taxon:32630"				
		/note="primer pAChE-Kpn, derived from human AChE gene and				


```

/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2"
/map="59C3-59C5"
/clone="DS02840 (D428)"
/clone_lib="P1 library, partial Sau3A in pAd10sacBII"

BASE COUNT      6308 a  5342 c  5216 g  6252 t  105 others
ORIGIN

Query Match      58.1%; Score 18; DB 2; Length 23223;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 gtagcgctgagcaatttg 31
|||||
Db 2921 GTAGCGCTGAGCAATTG 2938

RESULT 5
AC013153
LOCUS
DEFINITION
AC013153 146775 bp DNA linear HTG 03-NOV-1999
Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
AC013153
VERSION AC013153.1 GI:6223178
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 146775)
Adams,M. and Venter,J.C.
Direct Submission
Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM:10213761 by the submitter.
For further information on this sequence you may e-mail to
fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 146775
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT      4074 a  3285 c  3304 g  4010 t
ORIGIN

Query Match      58.1%; Score 18; DB 2; Length 146775;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 gtagcgctgagcaatttg 31
|||||
Db 57937 GTAGCGCTGAGCAATTG 57954

RESULT 6
AC099035
LOCUS
DEFINITION
AC099035 175451 bp DNA linear INV 08-NOV-2001
Drosophila melanogaster, chromosome 2R, region 59B-59C, BAC clone
BACR37H13, complete sequence.
AC099035
VERSION AC099035.1 GI:16798967
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
```

```

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 175451)
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattel,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirska,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 2R, region 59B-59C
Unpublished
2 (bases 1 to 175451)
Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattel,B., Moshrefi,A.,
McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
Stapleton,M., Strong,R., Svirska,R., Tector,C., Williams,S.M.,
Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Direct Submission
Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdgpe@fruitfly.berkeley.edu.

FEATURES
source
1. 175451
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="59B-59C"
/clone="BACR37H13 (D1380)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
PBACe3.6)"

BASE COUNT      48153 a  39588 c  39612 g  48098 t
ORIGIN

Query Match      58.1%; Score 18; DB 3; Length 175451;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 gtagcgctgagcaatttg 31
|||||
Db 21108 GTAGCGCTGAGCAATTG 21125

RESULT 7
AC008352/c
LOCUS
DEFINITION
AC008352 181463 bp DNA linear INV 27-APR-2001
Drosophila melanogaster, chromosome 2R, region 59C-59D, BAC clone
```

ACCESSION BACR08P09, complete sequence.
VERSION AC008352
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 181463)
AUTHORS Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome 2R, region 59C-59D
Unpublished
2 (bases 1 to 181463)
Celniker,S.E., Agbayani,A., Arcalna,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.
Direct Submission
TITLE Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
JOURNAL
COMMENT On Apr 27, 2001 this sequence version replaced gi:5670370.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdgp@fruitfly.berkeley.edu.
FEATURES
source Location/Qualifiers
1. 181463
/organism="Drosophila melanogaster"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="59C-59D"
/clone="BACR08P09 (D954)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
PBACE3.6)"
BASE COUNT 50024 a 40430 c 40416 g 50593 t
ORIGIN

Query Match 58.1%; Score 18; DB 3; Length 181463;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 gtagcgctgagcaatttg 31
|||||
Db 91011 GTAGCGCTGAGCAATTG 90994

RESULT 8
AE003459/c
LOCUS
DEFINITION Drosophila melanogaster genomic scaffold 142000013386038 section 8 of 15, complete sequence.
ACCESSION AE003459 AE002575
VERSION AE003459.2 GI:10727052
KEYWORDS HTG.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 299751)
AUTHORS Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S., Fleischmann,W., Fosler,C., Gabriellian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Helman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kimmel,B.E., Kodira,C.D., Kraft,C., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Lasko,P., Lei,Y., Levitsky,A.A., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spler,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
The genome sequence of Drosophila melanogaster
Science 287 (5461), 2185-2195 (2000)
20196006
2 (bases 1 to 299751)
ADAMS,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
TITLE Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
JOURNAL
COMMENT On Oct 9, 2000 this sequence version replaced gi:7291495.
FEATURES
source Location/Qualifiers
1. 299751
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2R"
complement(join(662..1075,1557..>1759))
/gene="CG3501"
/product="CT11797"
/db_xref="FLYBASE:FBan0003501"
/db_xref="FLYBASE:FBgn0034791"
complement(<662..>1759)
/gene="CG3501"

CDS

```
/map="59B2-59B3"
/db_xref="FLYBASE:FBan0003501"
/db_xref="FLYBASE:FBgn0034791"
complement(join(662. .1075,1557. .1754))
/gene="CG3501"
/note="CG3501 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003501"
/db_xref="FLYBASE:FBgn0034791"
/protein_id="AAF46921.1"
/db_xref="GI:7291496"
/translation="MCDYKVSERAYAKLIFHAAKYPHQAVNGLLLAEKTSKGSQVEIV
DAIPLFHQClyVTPMAEVALMLIDAHAREGLVIAGYYAAPENFYDNQVDKTPAAKIA
DKIQENFKNACFVVYDNKMLTLQHDRAAIQVFNCPGDSGARWSKAKFTLSQASDTLEG
VSLLLKRGAMRDVDFDNHLDNPDKNWTNDFLNQPLNDLQKLY"
complement(join(2131. .2980,3056. .3446,3503. .3816,
3875. .3948,4063. .4155,4209. .4398,4452. .4511,4569. .4966,
5071. .5175))
/gene="CG3499"
/product="CT11787"
/db_xref="FLYBASE:FBan0003499"
/db_xref="FLYBASE:FBgn0034792"
complement(<2131. .>5175)
/gene="CG3499"
```

mRNA

```
/map="59B2-59B3"
/db_xref="FLYBASE:FBan0003499"
/db_xref="FLYBASE:FBgn0034792"
complement(<2131. .>5175)
/gene="CG3499"
/product="CT11787"
/db_xref="FLYBASE:FBan0003499"
/db_xref="FLYBASE:FBgn0034792"
complement(join(2317. .2980,3056. .3446,3503. .3816,
3875. .3948,4063. .4155,4209. .4398,4452. .4511,4569. .4966,
5071. .5097))
/gene="CG3499"
/note="CG3499 gene product"
```

gene

CDS

```
/map="59B2-59B3"
/db_xref="FLYBASE:FBan0003499"
/db_xref="FLYBASE:FBgn0034792"
complement(join(2317. .2980,3056. .3446,3503. .3816,
3875. .3948,4063. .4155,4209. .4398,4452. .4511,4569. .4966,
5071. .5097))
/gene="CG3499"
/note="CG3499 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003499"
/db_xref="FLYBASE:FBgn0034792"
/protein_id="AAF46922.1"
/db_xref="GI:7291497"
/translation="MFSTTHSVPLYLGNFSRPHYYSVNRTKLHGSAGAARLSKST
STSSRSHDLVLDRLNLSRSSASIQGMVERAARLNGILDRLVDVVLAKVTSMLPSMR
DVRVTEESATQIGRVQLQNYQFEVSLTGAAGSVPTGANVKVIPTITPGLLRPLFSQQ
QLNQIRGFKTDRSIEAEQKRNPMTSRLKNALANSQORLDGDTPLQAEKLRLLAKSE
EHGFNKAESLKIAFAEGYLAAANSTSDSPKSGTKMYLKTIVIVVFLGIFLSFFTSTN
GSVFRIOLGNOQVEVDPEEINVTFFEDVKGCDEAKQELKEVVEFLKSPEKFSNLGGLPK
GVLLVGPPGTGKTLARAVAGEAKVPFFHAAGPEFDEVLVGQGARRVLDLPKAAKARA
PCVIFIDEIDSVGAKTNSVLPYANQTINQLLSEMDGFHQNAGVIVLGATNRRDDLD
QALLRPGRFDVEVMVSTPDFTGKKEILSLYTKILHDEIDLMDLARGTSGFTGADLEN
MINQAALRAAIDGAETVSMKHLETARDKVLMPGERKARLPDEEANTITAYHEGGHAIV
AFYTKESHPLHKVTIMRGPSPSLGHTAYIPEKERYHVTKAQLLAMDTMMGSRAAEELV
FGTDKITSGASSDLQATSIAATHMVRDWMGMSDKVGLRTIEASKGLGTGDTLGPNTIEA
VDAEIKRILSDSYERAKAILRKHTREHKALAEALLKYETLDADDIKAILNESQT"
join(<5798. .5857,5914. .6403,6465. .>6688)
/gene="CG13533"
/product="CT32908"
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
<5798. .>6688
/gene="CG13533"
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
join(5798. .5857,5914. .6403,6465. .6688)
/gene="CG13533"
/note="CG13533 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
/protein_id="AAF46923.1"
/db_xref="GI:7291498"
/translation="MDSPLNDGSHPPPHAPHPLADYQFSAEVKALRECNTEFFQF
SLPFGTGLGLLAYFGVKNGYLQGHVKYGAVPKVMGMVILGYFVGKFSYQQCAEKIMR
```

mRNA

```
/map="59B2-59B3"
/db_xref="FLYBASE:FBan0003501"
/db_xref="FLYBASE:FBgn0034791"
complement(join(662. .1075,1557. .1754))
/gene="CG3501"
/note="CG3501 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003501"
/db_xref="FLYBASE:FBgn0034791"
/protein_id="AAF46921.1"
/db_xref="GI:7291496"
/translation="MCDYKVSERAYAKLIFHAAKYPHQAVNGLLLAEKTSKGSQVEIV
DAIPLFHQClyVTPMAEVALMLIDAHAREGLVIAGYYAAPENFYDNQVDKTPAAKIA
DKIQENFKNACFVVYDNKMLTLQHDRAAIQVFNCPGDSGARWSKAKFTLSQASDTLEG
VSLLLKRGAMRDVDFDNHLDNPDKNWTNDFLNQPLNDLQKLY"
complement(join(2131. .2980,3056. .3446,3503. .3816,
3875. .3948,4063. .4155,4209. .4398,4452. .4511,4569. .4966,
5071. .5175))
/gene="CG3499"
/product="CT11787"
/db_xref="FLYBASE:FBan0003499"
/db_xref="FLYBASE:FBgn0034792"
complement(<2131. .>5175)
/gene="CG3499"
```

gene

```
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
<5798. .>6688
/gene="CG13533"
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
join(5798. .5857,5914. .6403,6465. .6688)
/gene="CG13533"
/note="CG13533 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
/protein_id="AAF46923.1"
/db_xref="GI:7291498"
/translation="MDSPLNDGSHPPPHAPHPLADYQFSAEVKALRECNTEFFQF
SLPFGTGLGLLAYFGVKNGYLQGHVKYGAVPKVMGMVILGYFVGKFSYQQCAEKIMR
```

CDS

```
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
join(5798. .5857,5914. .6403,6465. .6688)
/gene="CG13533"
/note="CG13533 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0013533"
/db_xref="FLYBASE:FBgn0034793"
/evidence=not_experimental
/protein_id="AAF46923.1"
/db_xref="GI:7291498"
/translation="MDSPLNDGSHPPPHAPHPLADYQFSAEVKALRECNTEFFQF
SLPFGTGLGLLAYFGVKNGYLQGHVKYGAVPKVMGMVILGYFVGKFSYQQCAEKIMR
```

mRNA

```
LPNSHLGELLRQRQGGVISSITPDENLGRAFTLAPFSPSSADVYSDEAYQGRSTS
LNLDTESRPTLSGLDDIYRPTLDSGSMLEAELPEPSKPGQSYEDLRRNRREYSKHQ
QSPYSRPEPPVAVQORPVEQAQSEPAGRKNQYGDSTWD"
complement(join(6781. .6948,7005. .7562,7621. .7735,
7790. .>7947))
/gene="CG3495"
/product="CT11779"
/db_xref="FLYBASE:FBan0003495"
/db_xref="FLYBASE:FBgn0034794"
complement(<6781. .>7947)
/gene="CG3495"
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0003495"
/db_xref="FLYBASE:FBgn0034794"
complement(join(6781. .6948,7005. .7562,7621. .7735,
7790. .7914))
/gene="CG3495"
/note="CG3495 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003495"
/db_xref="FLYBASE:FBgn0034794"
/protein_id="AAF46924.1"
/db_xref="GI:7291499"
/translation="MKKVLVTGGTGLVGKALEAVIKEQSPEDQWFFAGSKDADLTNL
AATQALFAREKPTHVHLAAMVGGLFHMNNNLDLNNLLINDNVLOTAEHQGCYKV
VSLSTCIFPDKTSYPIDETMVHNGPPHPSPNYGSYAKRLIDVQNHAYHDKYGRVYTS
VIPCNIFGPHDNYNPEVSHVIPGMIYRMHQLVTEKTDVPENDKVFVFGSGMPLRQFV
YSRDLAELMIWLVLRNYESVEPIILSADEVQEVTFEVAQAVAKAFNENGRVLCVDTSKS
DQYKKTASNAKLRSFLPDYAFDTLEAINASVKWYIENYDQARK"
join(<8169. .8207,8270. .8363,8418. .9382,9441. .9527,
9583. .9785,9838. .10214,10275. .10449,10506. .10994,
11053. .11379,11439. .11662,11721. .11803,11860. .12007,
12077. .12734,12796. .12936,12995. .>13304)
/gene="CG3695"
/product="CT12373"
/db_xref="FLYBASE:FBan0003695"
/db_xref="FLYBASE:FBgn0034795"
<8169. .>13304
/gene="CG3695"
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0003695"
/db_xref="FLYBASE:FBgn0034795"
join(8169. .8207,8270. .8363,8418. .9382,9441. .9527,
9583. .9785,9838. .10214,10275. .10449,10506. .10994,
11053. .11379,11439. .11662,11721. .11803,11860. .12007,
12077. .12734,12796. .12936,12995. .13304)
/gene="CG3695"
/note="CG3695 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003695"
/db_xref="FLYBASE:FBgn0034795"
/protein_id="AAF46925.1"
/db_xref="GI:7291500"
/translation="METQVIDTVNEFLKVDLSDEAFVSVIVFKPNTQEERATRFANDL
VTAFGNVAQENREQVLRLLRAAGASGYHIKVLMAALVKLVDAHVITARMLCDKVL
CEKLDHFHRTFWIESFRLIKRVIVQVDYKGVREINKVCRDKAQWFPLNVNVNVTYMPQLL
```

CDS

```
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0003495"
/db_xref="FLYBASE:FBgn0034794"
complement(join(6781. .6948,7005. .7562,7621. .7735,
7790. .7914))
/gene="CG3495"
/note="CG3495 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003495"
/db_xref="FLYBASE:FBgn0034794"
/protein_id="AAF46924.1"
/db_xref="GI:7291499"
/translation="MKKVLVTGGTGLVGKALEAVIKEQSPEDQWFFAGSKDADLTNL
AATQALFAREKPTHVHLAAMVGGLFHMNNNLDLNNLLINDNVLOTAEHQGCYKV
VSLSTCIFPDKTSYPIDETMVHNGPPHPSPNYGSYAKRLIDVQNHAYHDKYGRVYTS
VIPCNIFGPHDNYNPEVSHVIPGMIYRMHQLVTEKTDVPENDKVFVFGSGMPLRQFV
YSRDLAELMIWLVLRNYESVEPIILSADEVQEVTFEVAQAVAKAFNENGRVLCVDTSKS
DQYKKTASNAKLRSFLPDYAFDTLEAINASVKWYIENYDQARK"
join(<8169. .8207,8270. .8363,8418. .9382,9441. .9527,
9583. .9785,9838. .10214,10275. .10449,10506. .10994,
11053. .11379,11439. .11662,11721. .11803,11860. .12007,
12077. .12734,12796. .12936,12995. .>13304)
/gene="CG3695"
/product="CT12373"
/db_xref="FLYBASE:FBan0003695"
/db_xref="FLYBASE:FBgn0034795"
<8169. .>13304
/gene="CG3695"
/map="59B2-59B4"
/db_xref="FLYBASE:FBan0003695"
/db_xref="FLYBASE:FBgn0034795"
join(8169. .8207,8270. .8363,8418. .9382,9441. .9527,
9583. .9785,9838. .10214,10275. .10449,10506. .10994,
11053. .11379,11439. .11662,11721. .11803,11860. .12007,
12077. .12734,12796. .12936,12995. .13304)
/gene="CG3695"
/note="CG3695 gene product"
/codon_start=1
/db_xref="FLYBASE:FBan0003695"
/db_xref="FLYBASE:FBgn0034795"
/protein_id="AAF46925.1"
/db_xref="GI:7291500"
/translation="METQVIDTVNEFLKVDLSDEAFVSVIVFKPNTQEERATRFANDL
VTAFGNVAQENREQVLRLLRAAGASGYHIKVLMAALVKLVDAHVITARMLCDKVL
CEKLDHFHRTFWIESFRLIKRVIVQVDYKGVREINKVCRDKAQWFPLNVNVNVTYMPQLL
```

Query Match 58.1%; Score 18; DB 3; Length 299751;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 gtagcgctgagcaatttg 31
|||||
Db 171879 GTAGCGCTGAGCAATTG 171862

RESULT 9

AF099920 AF099920 34690 bp DNA linear INV 16-NOV-2001
LOCUS Caenorhabditis elegans cosmid H06H21, complete sequence.
DEFINITION
ACCESSION AF099920
VERSION AF099920.2 GI:13559632
KEYWORDS HTG.

SOURCE ORGANISM
Caenorhabditis elegans.
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 34690)
AUTHORS The C. elegans Sequencing Consortium.
TITLE Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium
JOURNAL Science 282 (5396), 2012-2018 (1998)
MEDLINE 99069613

REFERENCE 2 (bases 1 to 34690)
AUTHORS Bauer,C., Rohlfing,T. and Ahrens,C.
TITLE The sequence of C. elegans cosmid H06H21
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 34690)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 4 (bases 1 to 34690)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 34690)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, MO 63108, USA
REFERENCE 6 (bases 1 to 34690)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (06-APR-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 7 (bases 1 to 34690)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 8 (bases 1 to 34690)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 9 (bases 1 to 34690)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
COMMENT On Apr 6, 2001 this sequence version replaced gi:3811351.
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis , MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: rw@nematode.wustl.edu and jesus@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

NEIGHBORING COSMID INFORMATION

FEATURES source
1. .34690
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="IV"
/clone="H06H21"
complement(557..1048)
/gene="H06H21.6"
complement(join(557..820,874..984,1034..1048))
/gene="H06H21.6"
/note="coded for by the following C. elegans cDNAs:
yk61h5.3, yk81h5.3, yk109f12.3, yk199a4.3, yk233b11.5,
yk308d3.3, yk336f8.3, yk401e2.3, yk436b11.3"
/codon_start=1
/product="Hypothetical protein H06H21.6"
/protein_id="AAK29846.1"
/db_xref="GI:13559637"
/translation="MQVERMLTTRTKEMRQKDEQMTNRYKYTLIRVLPGNLLMOGV
FGCHEPFSAVRVFVASTLSDALSTSEFTLRDAASQLVEDESASLAQLSLAPALLHV
FAENLSDYGQIVADEHIEIIELEASD"
complement(2447..2931)
/gene="H06H21.7"
complement(join(2447..2469,2515..2662,2758..2931))
/gene="H06H21.7"
/note="coded for by the following C. elegans cDNAs:
yk109f12.5, yk199a4.5, yk676b9.5, yk401e2.5, yk436b11.5,
yk61h5.5, yk841e11.5"
/codon_start=1
/product="Hypothetical protein H06H21.7b"
/protein_id="AAL32226.1"
/db_xref="GI:16950432"
/translation="MGEPLVAEPTTEQTPLSTNSVAYQSSDALPDSSEAFGVENVYRA
LIIFCDNYSLDKKQVALGLFLSILLFIFSVGTIIAWNRYSPIIQEFIIRENIKARNK
SNEKNKKEE"
complement(join(2451..2464,2515..2662,2758..2931))
/gene="H06H21.7"
/note="coded for by the following C. elegans cDNAs:
yk81h5.5, yk122d1.5, yk308d3.5, yk737a1.5, yk676e6.5,
yk797c06.5, yk839d09.5, yk677f1.5"
/codon_start=1
/product="Hypothetical protein H06H21.7"
/protein_id="AAK29847.1"
/db_xref="GI:13559638"
/translation="MGEPLVAEPTTEQTPLSTNSVAYQSSDALPDSSEAFGVENVYRA
LIIFCDNYSLDKKQVALGLFLSILLFIFSVGTIIAWNRYSPIIQEFIIRENIKARNK
SNEKNKKEE"
3900..3901
/note="SL1 trans-splice site; see EST yk426b1.5"
3903..4647
/gene="H06H21.3"
join(3903..4002,4050..4334,4382..4647)
/gene="H06H21.3"
/note="similar to eukaryotic translation initiation
factors 1A (EIF-1A); coded for by the following C. elegans
cDNAs: cm13d5, yk109c11.3, yk109c11.5, yk176b3.5,
yk254a1.5, yk343d8.3, yk343d8.5, yk344e10.3, yk344e10.5,
yk371g1.3, yk371g1.5, yk415h10.3, yk415h10.5, yk426b1.3,
yk426b1.5, yk456f5.3, yk456f5.5, yk458b12.3, yk458b12.5,
yk474c1.3, yk474c1.5, yk484f2.5, yk497b8.3, yk497b8.5"
/codon_start=1
/product="Hypothetical protein H06H21.3"
/protein_id="AAK29845.1"
/db_xref="GI:13559636"
/translation="MPKNKGKGGKRRRGKNENDFMKRELDLKEEGQEYGVSKMLGN

gene	GRVQVFCFDGKQRVCHIRGKLRKKVWINVGDIIILVGLRDYQDDKGDVILKYTPDEARR LKNEGLIPENAKLNEDEQDEGEVEFLDHVCGEAGEAKSDSDSDSEKSDSEAG SDKEESDDDDSEEDSDSDNIREDLAAGRGFKEDTRRGGNRGGKNKYGKR"									
CDS	5018. .6996 /gene="H06H21.2" join(5018. .5220,5624. .5902,5950. .6076,6154. .6401, 6775. .6904,6955. .6996) /gene="H06H21.2" /codon_start=1 /product="Hypothetical protein H06H21.2" /protein_id="AAK29844.1" /db_xref="GI:13559635" /translation="MSQLAATCIHIVKIIYFSYFKLNLNEMIDRVLLYNLSTSAFPSY DNATRIDLFFELIAVEQVLAKIPEPPAPLSFFYMYWINVINDLFRRTSTWLSVMA TIRLIVLKESTRRAFRKVNFPKFGFTLVIGTLLSSIPLTVIYYFRDYIVKIGDWKPMN NCTFATSPESRSVIFTLVQSKVYQANDGLVGKVIQLINGTISKLPFCFLLPVLTMYLI VELKAKDHQRMNLNLTIRPVIINETSQSVASTERTTGLVVFITASTFLIEVPNGIVRVL QFGYTDLGYWRMATSVAQFSSAFVFLHAALQGAIFFLMSSQYRRRAVSKIFKNDRPPII IAVSSSYT" 7392. .10283 /gene="H06H21.1" join(7392. .7546,7602. .7820,8005. .8283,8329. .8455, 9474. .9727,10062. .10191,10242. .10283) /gene="H06H21.1" /codon_start=1 /product="Hypothetical protein H06H21.1" /protein_id="AAK29843.1" /db_xref="GI:13559634" /translation="MQVALIKNYSVQFGNSSQKKMDENTYNLATSAPPSYDNATKIQL YHFLKTVENFVHPSIALNIVFSIGVFTTSFHLVLSQKSILKSSVILIMMGVALCDI LAMVASVFYNILYLMIEFNVPKCEPPLSLPPFHMWINVINDFRRRCSTWLSVAMAL IRWIVMKYSTRRRAFRKVTLSKFGAYIFMGALALSPLTFIFYFRDYIVKIGDWKPLNN CSKSANFTVSLNIYNLVQSEVYTANDGLVGKVIQLINGTFSKLIPCFLLPILTTLLIA ELRKAQDHQRMNLNSTRLGNFPETFSQGVASTERTTGLVVFMTSTFVIEVPGGIVRV LQFGYTDLGYWRLATSVGQICSLFFVIHASLQGVIFFLMSSQYRRRTVSKIFKRDRPRV VIAVSSHFS" complement(10489. .14041) /gene="H06H21.8" complement(join(10489. .10683,10871. .11130,11944. .12067, 12543. .12851,13608. .13694,13850. .14041)) /gene="H06H21.8" /codon_start=1 /product="Hypothetical protein H06H21.8" /protein_id="AAK29842.1" /db_xref="GI:13559633" /translation="MKAKLKQLRLVSESSFPEIENFHEKVDASFEEKLENAKSFWSEIY VAHLKVVGDKVPESVFIKVPRISENVLRCEDESAVNHLNDVLLYYSKKENLFYKHF EYGSIPNFPFPKVIYFTEDINGEATGIVAEINLSEKVFAVEHIPGLKHEQILRLMEALA GLHSFLMKRDDKSYVESFVEGAHGRETFFSEGMQNMMFEEALTLENVSPEVFGNDRIRN									
Query Match 54.8%; Score 17; DB 3; Length 34690; Best Local Similarity 100.0%; Pred. No. 9.1; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
QY	1	cggtagctatcaggtag	17							
Db	26826	CGGTACCTATCAGGTAG	26842							
RESULT	10									
AC044838										
LOCUS	AC044838	178881 bp	DNA	linear	HTG 09-MAY-2001					
DEFINITION	Mus musculus chromosome 11 clone RP23-65B3 map 11,	WORKING DRAFT								
ACCESSION	AC044838	SEQUENCE, 6 unordered pieces.								
VERSION	AC044838.3	GI:12704668								
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.									
SOURCE	house mouse.									
ORGANISM	Mus musculus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
	1 (bases 1 to 178881)									
	Birren,B., Linton,L., Nusbaum,C. and Lander,E.									

TITLE JOURNAL REFERENCE AUTHORS	Mus musculus chromosome 11, clone RP23-65B3	
	Unpublished	
	2 (bases 1 to 178881)	
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,	
	Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,	
	Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,	
	Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,	
	Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,	
	Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,	
	Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,	
TITLE JOURNAL	Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,	
	Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,	
	Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczky,J.,	
	Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,	
	McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,	
	Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,	
	Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,	
	O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,	
	Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,	
	Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,	
COMMENT	Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,	
	Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,	
	Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,	
	Young,G., Zainoun,J., Zimmer,A. and Zody,M.	
	Direct Submission	
	Submitted (12-APR-2000) Whitehead Institute/MIT Center for Genome	
	Research, 320 Charles Street, Cambridge, MA 02141, USA	
	On Feb 7, 2001 this sequence version replaced gi:9369528.	
	All repeats were identified using RepeatMasker:	
	Smit, A.F.A. & Green, P. (1996-1997)	
TITLE JOURNAL	http://ftp.genome.washington.edu/RM/RepeatMasker.html	
	----- Genome Center	
	Center: Whitehead Institute/ MIT Center for Genome Research	
	Center code: WIBR	
	Web site: http://www-seq.wi.mit.edu	
	Contact: sequence_submissions@genome.wi.mit.edu	
	----- Project Information	
	Center project name: L8884	
	Center clone name: 65_B_3	
	----- Summary Statistics	
FEATURES source	Sequencing vector: M13; M77815; 40% of reads	
	Sequencing vector: Plasmid; n/a; 60% of reads	
	Chemistry: Dye-terminator Big Dye; 100% of reads	
	Assembly program: Phrap; version 0.960731	
	Consensus quality: 176862 bases at least Q40	
	Consensus quality: 177787 bases at least Q30	
	Consensus quality: 178109 bases at least Q20	
	Insert size: 169000; agarose-fp	
	Insert size: 178381; sum-of-contigs	
	Quality coverage: 11.5 in Q20 bases; agarose-fp	
FEATURES source	Quality coverage: 10.9 in Q20 b.	
	* NOTE: This is a 'working draft' sequence. It currently	
	* consists of 6 contigs. The true order of the pieces	
	* is not known and their order in this sequence record is	
	* arbitrary. Gaps between the contigs are represented as	
	* runs of N, but the exact sizes of the gaps are unknown.	
	* This record will be updated with the finished sequence	
	* as soon as it is available and the accession number will	
	* be preserved.	
	* 1 24550: contig of 24550 bp in length	
FEATURES source	* 24551 24650: gap of 100 bp	
	* 24651 39560: contig of 14910 bp in length	
	* 39561 39660: gap of 100 bp	
	* 39661 53388: contig of 13728 bp in length	
	* 53389 53488: gap of 100 bp	
	* 53489 85058: contig of 31570 bp in length	
	* 85059 85158: gap of 100 bp	
	* 85159 129894: contig of 44736 bp in length	
	* 129895 129994: gap of 100 bp	
	* 129995 178881: contig of 48887 bp in length.	
FEATURES source	Location/Qualifiers	
	1..178881	
FEATURES source	/organism="Mus musculus"	

```

/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-65B3"
/clone_lib="RPCI-23 Female Mouse BAC"
1. .24550
/clone="assembly_fragment"
clone_end:SP6
vector_side:left"
24651. .39560
/clone="assembly_fragment"
39661. .53388
/clone="assembly_fragment"
53489. .85058
/clone="assembly_fragment"
85159. .129894
/clone="assembly_fragment"
129995. .178881
/clone="assembly_fragment"
clone_end:T7
vector_side:right"
48194 a 40622 c 40173 g 49390 t 502 others
ORIGIN

Query Match      54.8%; Score 17; DB 2; Length 178881;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ggtacctatcaggtagc 18
|||||
Db      40278 GGTACCTATCAGGTAGC 40294

RESULT 11
AL606531      205165 bp      DNA      linear      HTG 30-JAN-2002
LOCUS
DEFINITION Mus musculus chromosome 11 clone RP23-104H11, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION AL606531
VERSION AL606531.6 GI:18135124
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Sims,S.
Direct Submission
Submitted (29-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 11, 2002 this sequence version replaced gi:17976609.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bm104H11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 204106 bases at least Q40
Consensus quality: 204442 bases at least Q30
Consensus quality: 204585 bases at least Q20
Insert size: 204665; sum-of-contigs
Insert size: 196989; 13.3% error; agarose-fp
Quality coverage: 8.71x in Q20 bases; sum-of-contigs Quality
coverage: 9.05x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
```

```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
FEATURES
    source      Location/Qualifiers
                1. .205165
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /chromosome="11"
                /clone="RP23-104H11"
                /clone_lib="RPCI-23"
                1. .16702
                /note="assembly_fragment:01528
                fragment_chain:1"
                clone_end:SP6
                vector_side:left"
                16803. .73634
                /note="assembly_fragment:04543
                fragment_chain:1"
                73735. .115816
                /note="assembly_fragment:00970
                fragment_chain:1"
                115917. .119281
                /note="assembly_fragment:03278
                fragment_chain:1"
                119382. .182835
                /note="assembly_fragment:04560
                fragment_chain:1"
                182936. .205165
                /note="assembly_fragment:01498
                fragment_chain:1"
                clone_end:T7
                vector_side:right"
BASE COUNT      55224 a 48879 c 47670 g 52892 t 500 others
ORIGIN

Query Match      54.8%; Score 17; DB 2; Length 205165;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ggtacctatcaggtagc 18
|||||
Db      145499 GGTACCTATCAGGTAGC 145515

RESULT 12
AC006719      209365 bp      DNA      linear      HTG 23-FEB-1999
LOCUS
DEFINITION Caenorhabditis elegans clone Y17G9, *** SEQUENCING IN PROGRESS ***,
2 unordered pieces.
ACCESSION AC006719
VERSION AC006719.1 GI:4263219
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 209365)
Waterston,R.H.
The sequence of Caenorhabditis elegans clone
Unpublished
2 (bases 1 to 209365)
Waterston,R.H.
Direct Submission
Submitted (23-FEB-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
```

* as soon as it is available and the accession number will
* be preserved.
* 1 3306: contig of 3306 bp in length
* 3307 3319: gap of unknown length
* 3320 209365: contig of 206046 bp in length.
FEATURES
 source
 Location/Qualifiers
 1..209365
 /organism="Caenorhabditis elegans"
 /db_xref="taxon:6239"
 /clone="Y17G9"
BASE COUNT 69867 a 35978 c 35861 g 67646 t 13 others
ORIGIN

Query Match 54.8%; Score 17; DB 2; Length 209365;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtagctatcaggtag 17
|||||
Db 117013 CGGTACCTATCAGGTAG 117029

RESULT 13
AL645602/c
LOCUS
DEFINITION
AL645602 271254 bp DNA linear HTG 30-JAN-2002
Mus musculus chromosome 11 clone RP23-79E13, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION
VERSION AL645602.4 GI:18476870
KEYWORDS
SOURCE HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Mclay, K.
Direct Submission
Submitted (24-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 1, 2002 this sequence version replaced gi:16973244.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bm79E13
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 266425 bases at least Q40
Consensus quality: 267646 bases at least Q30
Consensus quality: 268536 bases at least Q20
Insert size: 269954; sum-of-contigs
Insert size: 237088; 4.4% error; agarose-fp
Quality coverage: 7.15x in Q20 bases; sum-of-contigs Quality
coverage: 8.14x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
 Location/Qualifiers
 1..271254
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-79E13"
 /clone_lib="RPCI-23"
 1..10406
misc_feature

/note="assembly_fragment:03937
fragment_chain:1
clone_end:SP6
vector_side:left"
10507..37111
/note="assembly_fragment:05020
fragment_chain:1"
37212..54576
/note="assembly_fragment:03702
fragment_chain:1"
54677..60967
/note="assembly_fragment:01824
fragment_chain:1"
61068..67193
/note="assembly_fragment:02345
fragment_chain:2"
67294..86046
/note="assembly_fragment:04863
fragment_chain:2"
86147..133683
/note="assembly_fragment:01307
fragment_chain:2"
133784..150247
/note="assembly_fragment:04138
fragment_chain:3"
150348..172993
/note="assembly_fragment:03787
fragment_chain:3"
173094..181148
/note="assembly_fragment:01116"
219460..243872
/note="assembly_fragment:01623
fragment_chain:4"
243973..261235
/note="assembly_fragment:00717
fragment_chain:4"
261336..271254
/note="assembly_fragment:02417
fragment_chain:4
clone_end:T7
vector_side:right"
BASE COUNT 73218 a 63056 c 62264 g 71412 t 1304 others
ORIGIN

Query Match 54.8%; Score 17; DB 2; Length 271254;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtacctatcaggtagc 18
|||||
Db 239776 GGTACCTATCAGGTAGC 239760

RESULT 14
DSP276564
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
DSP276564 1019 bp rRNA linear BCT 19-NOV-2000
Desulfotomaculum sp. DEM-KMe99-1 16S rRNA gene.
AJ276564
AJ276564.1 GI:11230398
16S ribosomal RNA; 16S rRNA gene.
Desulfotomaculum sp. DEM-KMe99-1.
Desulfotomaculum sp. DEM-KMe99-1
Bacteria; Firmicutes; Bacillus/Clostridium group; Hellobacterium
group; Desulfotomaculum; environmental samples.
1 (bases 1 to 1019)
Stubner, S. and Meuser, K.
Detection of Desulfotomaculum in an Italian rice paddy soil by 16S
ribosomal nucleic acid analyses
FEMS Microbiol. Ecol. 34 (1), 73-80 (2000)
JOURNAL

PUBMED 11053738
REFERENCE 2 (bases 1 to 1019)
AUTHORS Stubner,S.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Stubner S., Biogeochemie,
Max-Planck-Institut f. terrestrische Mikrobiologie,
Karl-von-Frisch-Strasse, 35043 Marburg, GERMANY
FEATURES
source Location/Qualifiers
1..1019
/organism="Desulfotomaculum sp. DEM-KMe99-1"
/db_xref="taxon:118333"
/clone="DEM-KMe99-1"
/note="isolated from rice paddy soil"
1..1019
/gene="16S rRNA"
/product="16S ribosomal RNA"
1..1019
/gene="16S rRNA"
BASE COUNT 257 a 228 c 339 g 194 t 1 others
ORIGIN

Query Match 51.6%; Score 16; DB 1; Length 1019;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 caggtagcgctgagca 26
|||||
Db 986 CAGGTAGCGCTGAGCA 1001

RESULT 15
HSU34962/c
LOCUS HSU34962 1585 bp mRNA linear PRI 16-MAY-1996
DEFINITION Human transcription factor HCSX (hcsx) mRNA, complete cds.
ACCESSION U34962
VERSION U34962.1 GI:1314280
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1585)
AUTHORS Turbay,D., Wechsler,S.B., Blanchard,K.M. and Izumo,S.
TITLE Molecular cloning, chromosomal mapping, and characterization of the
human cardiac-specific homeobox gene hCsx
JOURNAL Mol. Med. 2 (1), 86-96 (1996)
MEDLINE 97056197
REFERENCE 2 (bases 1 to 1585)
AUTHORS Turbay,D.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-1995) David Turbay, Cardiovascular Research
Center, Department of Internal Medicine, The University of Michigan
Medical Center, 7220 MSRB III, 1150 West Medical Center Drive, Ann
Arbor, MI 48109, USA
FEATURES
source Location/Qualifiers
1..1585
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q35"
/clone="phCsx 1313"
/tissue_type="heart"
/clone_lib="lambda gt10 (Tamkun, et al., FASEB J. 5,
331-337)"
/dev_stage="adult"
1..1585
/gene="hCsx"
177..1151
/gene="hCsx"
/function="transcription factor"
/note="homeodomain containing protein"
/codon_start=1

/product="HCSX"
/protein_id="AAC50470.1"
/db_xref="GI:1314281"
/translation="MFSPALPTPFVKDILNLEQQQRSLAAAGELSARLEATLAPS
SCMLAAFKPEAYAGPEAAAPGLPRLRAELGRAPSPAKCASAFAFPAYFPAYSDPD
AKDPRAEKKELCALQKAVELEKTEADNAERPRARRRRKPRVLFSSQAQVYELERRFKQ
RYLSAPERDQLASVLKLTSTQVKIWFQNRRYKCKRQRQDTLELVGLPPPPPPARRI
AVPVLVRDGGKPCLGDSAPYAPAYGVLNPGYNAYPAYPGYGAACSPGYSCCTAAYPA
GPSPAQPATAAANNFVNFVGVDLNAVQSPGIPQSNVSTLHGIRAW"
misc_feature 588..768
/gene="hCsx"
/note="encodes helix-turn-helix motif"
BASE COUNT 256 a 569 c 497 g 263 t
ORIGIN

Query Match 51.6%; Score 16; DB 9; Length 1585;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 caggtagcgctgagca 26
|||||
Db 43 CAGGTAGCGCTGAGCA 28

Search completed: August 31, 2002, 22:35:31
Job time: 25160 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:51:27 ; Search time 824.51 Seconds
(without alignments)
64.553 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cgggtacctatcaggtagcgctgagcaatttg 31

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	31	AAS17546	Human acetylcholin
C 2	31	100.0	5767	AAS17547	Plasmid PTM034. S
C 3	31	100.0	14446	AAS17548	Plasmid PTM036. S
C 4	16	51.6	591	AAF08167	Fusarium venenatum
C 5	16	51.6	6751	AAS09962	Genomic DNA #2 enc
6	16	51.6	16075	AAV99811	Gum gene cluster.
7	16	51.6	16079	AAV70753	Sequence of segmen
8	15	48.4	971	AAI69821	Shiitake mushroom
C 9	15	48.4	1725	AAS17549	Synthetic human ac

C 10	15	48.4	1845	24	AAS17493	Human cDNA encodin
11	15	48.4	2050	22	AAI69822	Shiitake mushroom
12	15	48.4	2051	19	AAV38497	Nucleic acid of th
C 13	15	48.4	2089	24	ABI99790	Mouse ischaemic co
C 14	15	48.4	2253	11	AAQ05998	Sequence encoding
C 15	15	48.4	2256	16	AAQ99002	Human acetylcholin
16	15	48.4	6290	22	AAI69823	Shiitake mushroom
C 17	15	48.4	12113	24	AAS17492	Human acetylcholin
C 18	14	45.2	358	20	AAX41139	Human secreted pro
C 19	14	45.2	393	20	AAX40564	Human secreted pro
C 20	14	45.2	588	23	AAS74531	DNA encoding novel
C 21	14	45.2	900	19	AAV63933	Mycobacterium tube
C 22	14	45.2	900	20	AAX81040	Nucleotide sequenc
23	14	45.2	1105	21	AAC52511	Arabidopsis thalia
C 24	14	45.2	1269	23	AAS74534	DNA encoding novel
C 25	14	45.2	1707	23	AAS94180	DNA encoding novel
C 26	14	45.2	1776	23	AAS74532	DNA encoding novel
C 27	14	45.2	1907	19	AAV38355	Transcription fact
C 28	14	45.2	1907	21	AAZ36707	Nucleotide sequenc
C 29	14	45.2	1908	23	AAS94225	DNA encoding novel
C 30	14	45.2	1910	23	AAS74535	DNA encoding novel
C 31	14	45.2	1920	19	AAV40701	Human ELF3 protein
C 32	14	45.2	1996	21	AAC77813	Human cancer assoc
33	14	45.2	2112	23	ABL12431	Drosophila melanog
C 34	14	45.2	2353	23	AAS89049	DNA encoding novel
C 35	14	45.2	2379	23	ABL06673	Drosophila melanog
36	14	45.2	3774	23	AAS79055	DNA encoding novel
37	14	45.2	3774	23	AAS86684	DNA encoding novel
C 38	14	45.2	3774	23	AAS90257	DNA encoding novel
C 39	14	45.2	3776	23	AAS94279	DNA encoding novel
C 40	14	45.2	4233	23	ABL12430	Drosophila melanog
C 41	14	45.2	4543	23	ABL06672	Drosophila melanog
42	14	45.2	5679	23	ABL03552	Drosophila melanog
43	14	45.2	5823	23	AAS89361	DNA encoding novel
C 44	14	45.2	5828	23	ABL13644	Drosophila melanog
C 45	14	45.2	6048	19	AAV09029	Human hhl sodium c

ALIGNMENTS

RESULT 1

AAS17546
ID AAS17546 standard; DNA; 31 BP.

XX AAS17546;

XX 25-FEB-2002 (first entry)

DE Human acetylcholinesterase, PCR primer AChE-Kpn.

XX Human; acetylcholinesterase; AChE; PCR primer; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigamine bromide; ss.

OS Homo sapiens.

OS Synthetic.

XX WO200171014-A2.

XX 27-SEP-2001.

XX 16-MAR-2001; 2001WO-US08468.

XX 17-MAR-2000; 2000US-190440P.

PA (MORT/) MOR T.

PA (SORE/) SOREQ H.

PA (ARNT/) ARNTZEN C.

PA (MASO/) MASON H.

PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX Claim 11; Page 4; 42pp; English.
XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence represents PCR primer AChE-Kpn used to amplify
CC human AChE as described in the method of the invention.
XX Sequence 31 BP; 7 A; 7 C; 9 G; 8 T; 0 other;
SQ

Query Match 100.0%; Score 31; DB 24; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgggtacctatcaggtagcgctgagcaatttg 31
Db 1 cgggtacctatcaggtagcgctgagcaatttg 31

RESULT 2
AAS17547/c
ID AAS17547 standard; DNA; 5767 BP.
XX
AC AAS17547;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM034.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.

PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX Claim 11; Page 28-31; 42pp; English.
XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM034, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;
SQ

Query Match 100.0%; Score 31; DB 24; Length 5767;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgggtacctatcaggtagcgctgagcaatttg 31
Db 2571 CGGTACCTATCAGGTAGCGCTGAGCAATTG 2541

RESULT 3
AAS17548/c
ID AAS17548 standard; DNA; 14446 BP.
XX
AC AAS17548;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM036.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM036;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.

```
XX 16-MAR-2001; 2001WO-US08468.
XX
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX WPI; 2002-055120/07.
DR
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
XX Claim 11; Page 32-41; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM036, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;
```

```
XX Fusarium venenatum.
OS
XX WO200056762-A2.
PN
XX
XX 28-SEP-2000.
PD
XX
XX 22-MAR-2000; 2000WO-US07781.
PF
XX
XX 22-MAR-1999; 99US-0273623.
PR
XX
XX (NOVO ) NOVO NORDISK BIOTECH INC.
PA (NOVO ) NOVO NORDISK AS.
PA
XX
PI Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX WPI; 2000-594572/56.
DR
XX Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
XX Claim 86; Page 648; 3161pp; English.
PS
XX
CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC all specifically claimed in the present invention.
XX
SQ Sequence 591 BP; 150 A; 154 C; 167 G; 119 T; 1 other;
```

```
Query Match 51.6%; Score 16; DB 21; Length 591;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 tacctatcaggtagcg 19
| | | | | | | | | | | | | | |
Db 474 TACCTATCAGGTAGCG 459

RESULT 5
AAS09962/c
ID AAS09962 standard; DNA; 6751 BP.
XX
AC AAS09962;
XX
DT 24-OCT-2001 (first entry)
XX
DE Genomic DNA #2 encoding human Csx/Nkx2.5.
XX Csx/Nkx2.5; cardiac enhancer; cardiac cell; cardiomyocyte induction;
KW therapeutic; heart tissue; gene therapy; human; ds.
KW
```


CC codes for protein products. Each gene is designated by a letter (see Fig 11) and its protein product is designated by that letter

CC preceded by "gp" (AAP70455-67).

XX

SQ Sequence 16079 BP; 2707 A; 4649 C; 5261 G; 3462 T; 0 other;

Query Match 51.6%; Score 16; DB 8; Length 16079;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 tcaggtagcgctgagc 25
|||||
Db 11355 tcaggtagcgctgagc 11370

RESULT 8
AAI69821
ID AAI69821 standard; DNA; 971 BP.
XX
AC AAI69821;
XX
DT 18-DEC-2001 (first entry)
XX
DE Shiitake mushroom tyrosinase gene terminator sequence.
XX
KW Shiitake mushroom; tyrosinase; promoter; terminator;
KW transcription control; ds.
XX
OS Lentinula edodes.
XX
PN JP2001157586-A.
XX
PD 12-JUN-2001.
XX
PF 01-DEC-1999; 99JP-0342347.
XX
PR 01-DEC-1999; 99JP-0342347.
XX
PA (IWAT-) IWATE KEN.
XX
DR WPI; 2001-599774/68.
XX
PT New DNA which can function as a promoter for ordering the initiation of
PT transcription in the tyrosinase gene of Lentinula edodes -

XX Claim 3; Page 13-14; 28pp; Japanese.
XX
CC The invention relates to a DNA sequences that can be used to
CC control the initiation of transcription of the tyrosinase gene of
CC Lentinula edodes. The promoter region and terminator region of the
CC Lentinula edodes tyrosinase gene were isolated. The DNA that can
CC function as a promoter comprises a 2639 bp sequence given in the
CC specification, or a sequence in which a base is deleted, replaced or
CC added with respect to the 2639 bp sequence. The DNA that can function
CC as a terminator comprises a 971 bp sequence given in the specification,
CC or a sequence in which a base is deleted, replaced or added with respect
CC to the 971 base pair sequence. The present sequence is a terminator
CC region of the invention.
XX
SQ Sequence 971 BP; 274 A; 204 C; 245 G; 248 T; 0 other;

Query Match 48.4%; Score 15; DB 22; Length 971;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 caggtagcgctgagc 25
|||||
Db 83 caggtagcgctgagc 97

RESULT 9

AAS17549/c

ID AAS17549 standard; DNA; 1725 BP.
XX
AC AAS17549;
XX
DT 25-FEB-2002 (first entry)
XX
DE Synthetic human acetylcholinesterase gene.
XX
KW Human; acetylcholinesterase; AChE; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX
DR WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -

Claim 11; Page 41-42; 42pp; English.

The invention relates to a method of producing a transgenic plant which contains a polynucleotide that encodes human acetylcholinesterase (AChE) which upon purification is effective against acetylcholinesterase poisoning. The method is used for treating a victim of acetylcholinesterase poisoning by administering a therapeutic amount of a physiologically active human acetylcholinesterase expressed in plant tissue. The extensive use of anticholinesterase pesticides with concurrent accidental poisoning, the threat of chemical warfare and environmental concerns demand the development of effective, inexpensive and stage countermeasures and bioremediation solutions. Prior art methods for treating AChE poisoning have used the muscarinic receptor antagonist atropine and oximes to reactivate the organophosphate(OP)-modified AChE. The reversible carbamate, pyridostigmine bromide has also been used as a prophylactic. However, these conventional treatments have limited effectiveness and serious short and long-term side effects and may result in significant performance deficits and even permanent brain damage. This invention permits the utilisation of cholinesterases to counter-act the toxic effects of anti-cholinergic agents. Using transgenic plants for the production of the enzymes is cost effective and the product is stable so that the injected enzymes have the advantage of having a long half-life. The transgenic form of the enzymes are also easy to purify. The present sequence represents synthetic human AChE gene optimised for expression in plants.

Sequence 1725 BP; 273 A; 553 C; 571 G; 328 T; 0 other;

Query Match 48.4%; Score 15; DB 24; Length 1725;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31

Db 1718 GCGCTGAGCAATTG 1704

RESULT 10

AA17493/c

ID AAS17493 standard; cDNA; 1845 BP.

XX AAS17493;

AC AAS17493;

XX 26-FEB-2002 (first entry)

XX Human cDNA encoding acetylcholinesterase, ACHE.

DE Human; ss; ACHE; acetylcholinesterase; polymorphic variant; haplotyping;

XX genotyping; neurological disease; Parkinson's disease;

KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.

XX Homo sapiens.

OS

XX

FF Key Location/Qualifiers

FT CDS 1..1845

FT /tag= a

FT /product= "ACHE"

FT replace (36,T)

FT /tag= b

FT /standard_name= "Single nucleotide polymorphism"

FT replace (101,A)

FT /tag= c

FT /standard_name= "Single nucleotide polymorphism"

FT replace (183,C)

FT /tag= d

FT /standard_name= "Single nucleotide polymorphism"

FT replace (838,G)

FT /tag= e

FT /standard_name= "Single nucleotide polymorphism"

FT replace (873,T)

FT /tag= f

FT /standard_name= "Single nucleotide polymorphism"

FT replace (1032,A)

FT /tag= g

FT /standard_name= "Single nucleotide polymorphism"

FT replace (1057,A)

FT /tag= h

FT /standard_name= "Single nucleotide polymorphism"

FT replace (1431,T)

FT /tag= i

FT /standard_name= "Single nucleotide polymorphism"

XX WO200179219-A2.

PN

XX 25-OCT-2001.

PD

XX 11-APR-2001; 2001WO-US11853.

PF

XX 14-APR-2000; 2000US-197173P.

PR

XX (GENA-) GENAISSANCE PHARM INC.

PA (KAZE/) KAZEMI A.

XX Bentivegna SC, Chew A, Choi JY, Koshy B;

PI

XX WPI; 2002-055248/07.

DR P-PSDB; AAU11231.

DR

XX New polymorphic variants comprising acetylcholinesterase (ACHE)

PT isogene, useful in expressing ACHE protein for use in screening for

PT candidate drugs to treat diseases related to ACHE activity, e.g.

PT neurological diseases or cancer -

XX

PS Claim 26; Fig 2; 79pp; English.

XX The invention relates to a polynucleotide comprising a polymorphic

CC

CC variant of an acetylcholinesterase (ACHE) gene or fragment, protein or

CC complement, the variant comprising an ACHE isogene defined by a haplotype

CC selected from haplotypes 1-20 listed in the specification. Also included

CC are methods for haplotyping and genotyping the ACHE gene of an

CC individual, a method for predicting a haplotype pair for the ACHE gene of

CC an individual, a method for identifying an association between a trait

CC and at least one haplotype or haplotype pair of ACHE gene, recombinant

CC nonhuman organisms transformed or transfected with the polynucleotide

CC where the organism expresses ACHE protein encoded by the first

CC nucleotide sequence or encoded by the polymorphic variant sequence,

CC an isolated antibody specific for and immunoreactive with ACHE,

CC a method of screening for drugs targeting the polypeptide contacting ACHE

CC polymorphic variant with a candidate agent and assaying for binding

CC activity, a computer system for storing and analysing polymorphism data

CC for ACHE gene and a genome anthology for ACHE gene which comprises ACHE

CC isogenes defined by haplotypes 1-20 given in the specification.

CC The Polymorphisms are useful for studying the biological function of

CC ACHE as well as in identifying drugs targeting this protein for the

CC treatment of disorder related to its abnormal expression or function.

CC The polymorphic variants may also be used in screening for compounds

CC targeting ACHE to treat a specific condition or disease predicted to be

CC associated with ACHE activity e.g. neurological diseases (e.g.

CC Parkinson's disease and Alzheimer's disease), cancer, leukaemia,

CC and tumours. The ACHE gene maps to human chromosome 7q22. The

CC present sequence is the coding sequence of the ACHE gene.

XX

SQ Sequence 1845 BP; 302 A; 590 C; 608 G; 345 T; 0 other;

Query Match 48.4%; Score 15; DB 24; Length 1845;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31

Db 1718 GCGCTGAGCAATTG 1704

RESULT 11

AA169822

ID AAI69822 standard; DNA; 2050 BP.

XX

AC AAI69822;

XX

DT 18-DEC-2001 (first entry)

XX

DE Shiitake mushroom tyrosinase gene.

XX Shiitake mushroom; tyrosinase; promoter; terminator;

KW transcription control; ds.

KW

XX Lentinula edodes.

OS

XX JP2001157586-A.

PN

XX 12-JUN-2001.

PD

XX 01-DEC-1999; 99JP-0342347.

PF

XX 01-DEC-1999; 99JP-0342347.

PR

XX (IWAT-) IWATE KEN.

PA

XX WPI; 2001-599774/68.

DR P-PSDB; AAG67089.

DR

XX New DNA which can function as a promoter for ordering the initiation of

PT transcription in the tyrosinase gene of Lentinula edodes -

PT

XX Example 1; Page 14-17; 28pp; Japanese.

PS

XX The invention relates to a DNA sequences that can be used to

CC control the initiation of transcription of the tyrosinase gene of

CC Lentinula edodes. The promoter region and terminator region of the
CC Lentinula edodes tyrosinase gene were isolated. The DNA that can
CC function as a promoter comprises a 2639 bp sequence given in the
CC specification, or a sequence in which a base is deleted, replaced or
CC added with respect to the 2639 bp sequence. The DNA that can function
CC as a terminator comprises a 971 bp sequence given in the specification,
CC or a sequence in which a base is deleted, replaced or added with respect
CC to the 971 base pair sequence. The present sequence is the
CC tyrosinase gene of Lentinula edodes.
XX
SQ Sequence 2050 BP; 559 A; 509 C; 469 G; 513 T; 0 other;

Query Match 48.4%; Score 15; DB 22; Length 2050;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 caggtagcgctgagc 25
|||||
Db 1974 caggtagcgctgagc 1988

RESULT 12
AAV38497
ID AAV38497 standard; cDNA to mRNA; 2051 BP.
XX
AC AAV38497;
XX
DT 17-SEP-1998 (first entry)
XX
DE Nucleic acid of the specification.
XX
KW Shiitake tyrosinase protein; Hokken number 57; ds.
XX
OS Lentinus edodes.
XX
FH Key Location/Qualifiers
FT CDS 21..1904
FT /*tag= a
XX
PN JP10174586-A.
XX
PD 30-JUN-1998.
XX
PF 17-DEC-1996; 96JP-0337127.
XX
PR 17-DEC-1996; 96JP-0337127.
XX
PA (IWAT-) IWATE KEN.
XX
DR WPI; 1998-421168/36.
DR N-PSDB; AAW62553.
XX
PT Shiitake tyrosinase gene - for producing large amounts of tyrosinase
PT when introduced into various cells
XX
PS Disclosure; Pages 7-10; 10pp; Japanese.
XX
CC The present sequence appears in the specification, which describes
CC a Shiitake tyrosinase protein. The nucleic acid sequence was isolated
CC from Shiitake strain Hokken number 57. The gene can produce a large
CC amount of tyrosinase when introduced to various cells.
XX
SQ Sequence 2051 BP; 559 A; 512 C; 469 G; 511 T; 0 other;

Query Match 48.4%; Score 15; DB 19; Length 2051;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 caggtagcgctgagc 25
|||||
Db 1975 caggtagcgctgagc 1989

RESULT 13
ABI99790/c
ID ABI99790 standard; cDNA; 2089 BP.
XX
AC ABI99790;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:890.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2; Page 2244-2245; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 2089 BP; 409 A; 638 C; 596 G; 446 T; 0 other;

Query Match 48.4%; Score 15; DB 24; Length 2089;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||
Db 1718 GCGCTGAGCAATTG 1704

RESULT 14
AAQ05998/c
ID AAQ05998 standard; DNA; 2253 BP.
XX
AC AAQ05998;
XX
DT 16-JAN-1991 (first entry)
XX

DE Sequence encoding human acetylcholinesterase (hAChE).
XX
KW Organophosphorous poisoning; OP; cancer; leukaemia;
KW megakaryocytopoiesis; ovarian cancer; ds.
XX
OS Homo sapiens.
XX
PN EP388906-A.
XX
PD 26-SEP-1990.
XX
XX
PF 20-MAR-1990; 90EP-0105274.
XX
PR 21-MAR-1989; 89IL-0089703.
XX
PA (YISS) YISSUM RES DEV CO.
XX
PI Soreq H, Zakut H;
XX
DR WPI; 1990-291865/39.
DR P-PSDB; AAR06989.
XX
PT Human acetyl:cholinesterase DNA and prodn. of recombinant hAChE -
PT for treatment of organo-phosphorous poisoning and diagnosis of haemo-
PT cytopoeitic disorders and ovarian carcinomas.
XX
PS Claim 5; Page 25; 47pp; English.
XX
CC Gene product is useful as an active pharmacological component for the
CC prophylaxis and treatment of organophosphorous poisoning, and post-
CC surgical apnea due to succinylcholine administration.
CC cDNA probe to the sequence may be used in diagnosis of various
CC leukaemias, abnormal megakaryocytopoiesis and ovarian carcinomas.
XX
SQ Sequence 2253 BP; 390 A; 740 C; 694 G; 429 T; 0 other;

Query Match 48.4%; Score 15; DB 11; Length 2253;
Best Local Similarity 100.0%; Pred. NO. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
| | | | | | | | | |
Db 1874 GCGCTGAGCAATTG 1860

RESULT 15
AAQ99002/c
ID AAQ99002 standard; DNA; 2256 BP.
XX
AC AAQ99002;
XX
DT 31-MAR-1996 (first entry)
XX
DE Human acetylcholinesterase (AChE) gene.
XX
KW Acetylcholinesterase; acetyl cholinesterase; EC-3.1.1.7;
KW chromosome-7q22; acetylcholine-hydrolyzing enzyme; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 160..2206
FT /*tag= a
XX
PN WO9523158-A1.
XX
PD 31-AUG-1995.
XX
PF 28-FEB-1995; 95WO-US02806.
XX
PR 09-JAN-1995; 95US-0370156.
PR 28-FEB-1994; 94US-0202755.

XX
PA (KOHN/) KOHN K I.
PA (YISS) YISSUM RES & DEV CO.
XX
PI Shani M, Soreq H, Zakut H;
XX
DR WPI; 1995-311499/40.
DR P-PSDB; AAR80726.
XX
PT Alternative forms of human acetyl cholinesterase (ChE) gene -
PT expressed in transgenic animal assay system for evaluating anti-ChE
PT activity of organo:phosphate(s), etc. or as model of ChE imbalance
XX
PS Claim 3; Fig.1A; 55pp; English.
XX
CC This DNA sequence encoding human acetylcholinesterase is useful
CC for producing transgenic animals which express AChE. The
CC transgenic animals are in turn useful as an assay system for
CC determining the anti-ChE activity of organophosphates, carbamates,
CC anti-ChE drugs, plant glycoalkaloids and snake venoms. This gene
CC contains a promoter region, 6 exons (E1 to E6) and 4 introns (I1-
CC I4).
XX
SQ Sequence 2256 BP; 390 A; 757 C; 680 G; 429 T; 0 other;

Query Match 48.4%; Score 15; DB 16; Length 2256;
Best Local Similarity 100.0%; Pred. NO. 19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
| | | | | | | | | |
Db 1877 GCGCTGAGCAATTG 1863

Search completed: August 31, 2002, 22:51:31
Job time: 16392 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 19:58:44 ; Search time 6064.54 Seconds
(without alignments)
68.992 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtaacctatcaggtagcgctgagcaatttg 31

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16	51.6	547	12 AZ586423	AZ586423 1M0392M15
C 2	16	51.6	694	12 AZ274865	AZ274865 RPCI-23-1
C 3	16	51.6	976	12 CNS02V61	AL215362 Tetraodon
C 4	16	51.6	1001	10 BI596784	BI596784 603242870
C 5	15	48.4	52	9 AI205485	AI205485 qr26b01.x
C 6	15	48.4	148	9 AW418501	AW418501 ha19d06.x
C 7	15	48.4	177	10 BF892364	BF892364 CM1-MT010
C 8	15	48.4	203	9 AW805755	AW805755 QV1-UM010
C 9	15	48.4	214	10 BF875285	BF875285 QV3-ET009
C 10	15	48.4	254	10 T28280	T28280 EST35080 Hu
C 11	15	48.4	260	9 AW138046	AW138046 UI-H-B11-
C 12	15	48.4	292	9 AI205484	AI205484 qr26a12.x
C 13	15	48.4	302	9 AA346027	AA346027 EST52141
C 14	15	48.4	348	9 BE146509	BE146509 RCI-HT021
C 15	15	48.4	352	10 R28476	R28476 yh50g12.r1
C 16	15	48.4	353	9 AI651502	AI651502 wb06e06.x
C 17	15	48.4	353	9 AI652210	AI652210 wb20c05.x

C 18	15	48.4	353	9 AW590331	AW590331 hg43c09.x
C 19	15	48.4	359	10 N27052	N27052 yx19e08.r1
C 20	15	48.4	367	9 AI341526	AI341526 qx92c12.x
C 21	15	48.4	367	10 BE502360	BE502360 hy22e02.x
C 22	15	48.4	378	10 F27586	F27586 HSPD15598 H
C 23	15	48.4	380	9 AA933814	AA933814 ol84b09.s
C 24	15	48.4	384	9 AI632562	AI632562 wb10e06.x
C 25	15	48.4	384	9 AW612795	AW612795 hh33a07.x
C 26	15	48.4	384	10 BE502241	BE502241 hyl4g01.x
C 27	15	48.4	384	10 BE504052	BE504052 hz34e12.x
C 28	15	48.4	393	9 AA998511	AA998511 UI-R-C0-i
C 29	15	48.4	395	9 AI766812	AI766812 w189a12.x
C 30	15	48.4	395	9 AI847791	AI847791 UI-M-AK1-
C 31	15	48.4	407	10 BE551703	BE551703 hx98f02.x
C 32	15	48.4	410	9 AI360141	AI360141 qv83d10.x
C 33	15	48.4	411	9 AI690171	AI690171 tx33a02.x
C 34	15	48.4	413	9 AA401251	AA401251 zv66b01.s
C 35	15	48.4	420	9 AI825097	AI825097 wb16a08.x
C 36	15	48.4	434	10 BI131084	BI131084 G114P89Y
C 37	15	48.4	442	9 BB733073	BB733073 BB733073
C 38	15	48.4	442	10 BG381540	BG381540 UI-R-CT0-
C 39	15	48.4	443	10 BG381528	BG381528 UI-R-CT0-
C 40	15	48.4	455	9 AW871154	AW871154 ra55e04.y
C 41	15	48.4	467	9 AW828558	AW828558 ra71c01.y
C 42	15	48.4	469	9 AI831696	AI831696 wj09e05.x
C 43	15	48.4	476	12 AQ661569	AQ661569 HS_5192_A
C 44	15	48.4	479	9 AW783232	AW783232 ra26c05.y
C 45	15	48.4	480	10 N36355	N36355 yx82f03.r1

ALIGNMENTS

RESULT 1
AZ586423/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

AZ586423 547 bp DNA linear GSS 13-DEC-2000
1M0392M15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0392M15 F, DNA sequence.
AZ586423 GI:11708613
GSS.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 547)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Unlversity of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0392 row: M column: 15
Seq primer: CGTTGTAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 547.
Location/Qualifiers
1. .547
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0392M15"
/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g1|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 179 a 130 c 122 g 116 t
ORIGIN

Query Match 51.6%; Score 16; DB 12; Length 547;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 aggtagcgctgagcaa 27
|||||
Db 304 AGGTAGCGCTGAGCAA 289

RESULT 2
AZ274865/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ274865
RPCI-23-118H10-TV RPCI-23 Mus musculus genomic clone RPCI-23-118H10
DNA sequence.
AZ274865
AZ274865.1 GI:9488482
GSS.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 694)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-118H10.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 118 row: H column: 10
Seq primer: T7
Class: BAC ends.

FEATURES
source
1. .694
/organism="Mus musculus"
/strain="C57BL/6J"

/db_xref="taxon:10090"
/clone="RPCI-23-118H10"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
ECORI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
ECORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 237 a 90 c 94 g 273 t
ORIGIN

Query Match 51.6%; Score 16; DB 12; Length 694;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gtacctatcaggtagc 18
|||||
Db 146 GTACCTATCAGGTAGC 131

RESULT 3
CNS02V61
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

CNS02V61
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
173P07 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL215362
AL215362.1 GI:7874181
GSS; genome survey sequence.
Tetraodon nigroviridis.
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 976)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 976)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 976)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.
Location/Qualifiers
1. .976
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="173P07"
/clone_lib="G"
/note="Genoscope sequence ID : COAG173CH04SP1-end :
PUC-Ori"

BASE COUNT 267 a 194 c 216 g 272 t 27 others
ORIGIN

Query Match 51.6%; Score 16; DB 12; Length 976;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 tagcgctgagcaattt 30
|||||
Db 670 TAGCGCTGAGCAATT 685

RESULT 4
BI596784/c
LOCUS
DEFINITION BI596784 1001 bp mRNA linear EST 07-SEP-2001
603242870F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285512 5',
mRNA sequence.

ACCESSION BI596784
VERSION BI596784.1 GI:15489723
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1001)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11721 row: k column: 17
High quality sequence stop: 67.

FEATURES
source
1..1001
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285512"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 169 a 296 c 309 g 227 t
ORIGIN

Query Match 51.6%; Score 16; DB 10; Length 1001;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ggtagcgctgagcaat 28
|||||
Db 465 GGTAGCGCTGAGCAAT 450

RESULT 5
AI205485/c
LOCUS
DEFINITION qr26b01.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:1941961 3',
similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);, mRNA

sequence.
ACCESSION AI205485
VERSION AI205485.1 GI:3764157
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 52)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES
source
1..52
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1941961"
/clone_lib="NCI_CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 8 a 23 c 11 g 10 t
ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ggcgtgagcaatttg 31
|||||
Db 47 GCGCTGAGCAATTG 33

RESULT 6
AW418501/c
LOCUS
DEFINITION AW418501 148 bp mRNA linear EST 09-FEB-2000
ha19d06.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2874155 3',
similar to TR:Q16169 Q16169 ACETYLCHOLINESTERASE ;, mRNA sequence.

ACCESSION AW418501
VERSION AW418501.1 GI:6946383
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 148)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL
COMMENT
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES
source

1. .148
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2874155"
/clone_lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 17 a 72 c 31 g 28 t
ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||
Db 15 GCGCTGAGCAATTG 1

RESULT 7

BF892364
LOCUS
DEFINITION CM1-MT0102-041100-434-d07 MT0102 Homo sapiens cDNA, mRNA sequence. EST 18-JAN-2001
ACCESSION BF892364
VERSION BF892364.1 GI:12283823
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 177)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

REFERENCE
AUTHORS

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM1&t2=CM1-MT0102-041100-434-d07&t3=2000-11-04&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 5
High quality sequence stop: 177.

FEATURES
source

1. .177
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0102"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 ,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 39 a 54 c 41 g 43 t
ORIGIN

Query Match 48.4%; Score 15; DB 10; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cggtacctatcaggt 15
|||||
Db 61 CGGTACCTATCAGGT 75

RESULT 8

AW805755
LOCUS
DEFINITION QV1-UM0105-180400-162-b06 UM0105 Homo sapiens cDNA, mRNA sequence. EST 16-MAY-2000
ACCESSION AW805755
VERSION AW805755.1 GI:7857625
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 203)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

REFERENCE
AUTHORS

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV1&t2=QV1-UM0105-180400-162-b06&t3=2000-04-18&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 35
High quality sequence stop: 203.
FEATURES
source
1. .203
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UM0105"
/dev_stage="Adult"
/note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 47 a 34 c 47 g 75 t
ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtacctatcaggta 16
|||||
Db 43 GGTACCTATCAGGTA 57

RESULT 9
BF875285
LOCUS BF875285 214 bp mRNA linear EST 17-JAN-2001
DEFINITION QV3-ET0099-101100-390-e10 ET0099 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF875285
VERSION BF875285.1 GI:12265415
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV3&t2-QV3-ET0099-
101100-390-e10&t3-2000-11-10&t4-1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 213.
FEATURES
source
1. .214
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0099"
/dev_stage="Adult"
/note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 48 a 34 c 45 g 87 t
ORIGIN

Query Match 48.4%; Score 15; DB 10; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 tacctatcaggtagc 18
|||||
Db 16 TACCTATCAGGTAGC 30

RESULT 10
T28280/c
LOCUS T28280 254 bp mRNA linear EST 06-SEP-1995
DEFINITION EST35080 Human Embryo Homo sapiens cDNA 5' end similar to
acetylcholinesterase (HT:518), mRNA sequence.
ACCESSION T28280
VERSION T28280.1 GI:610378
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A.,
Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D., FitzGerald
,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M., Glodek,A.,
Gnehm,C.L., Hanna,M.C., Hedblom,E., HinkleJr,P.S., Kelley,J.M.,
Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J.,
Dimke,D., Feng,P., Ferrie,A., Fischer,C., Hastings,G.A., He,W.-W.,
Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L.,
Kunsch,C., Ji,H., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei
,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon
,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
Venter,J.C.

TITLE Initial Assessment of Human Gene Diversity and Expression Patterns
Based Upon 83 Million Basepairs of cDNA Sequence
JOURNAL Nature 377, 3-174 (1995)
MEDLINE 96026280
COMMENT Other_ESTs: THC20776
Contact: Venter, JC
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: tdbinfo@tdb.tigr.org
For clone availability, additional sequence and expression
information related to this EST, please contact the TIGR Database
(tdbinfo@tdb.tigr.org)
Seq primer: M13 Reverse.
Location/Qualifiers
1. .254
/organism="Homo sapiens"
/db_xref="ATCC (inhost):101797"
/db_xref="taxon:9606"
/clone_lib="Human Embryo"
/tissue_type="embryo"
BASE COUNT 45 a 89 c 78 g 41 t 1 others
ORIGIN

FEATURES
source
1. .214
/organism="Homo sapiens"
/db_xref="ATCC (inhost):101797"
/db_xref="taxon:9606"
/clone_lib="Human Embryo"
/tissue_type="embryo"
BASE COUNT 45 a 89 c 78 g 41 t 1 others
ORIGIN

```

Query Match      48.4%; Score 15; DB 10; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||
Db 101 GCGCTGAGCAATTG 87

RESULT 11
AW138046/c
LOCUS
DEFINITION
  UI-H-B11-abw-b-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
  IMAGE:2713294 3', mRNA sequence.
ACCESSION
  AW138046
VERSION
  AW138046.1 GI:6142364
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 260)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Oligo-dt track not found, Not 1 site shown in beginning of sequence
  is likely internal to the message. CDNA Library Preparation: M.B.
  Soares Lab Clone distribution: NCI-CGAP clone distribution
  information can be found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html The following repetitive
  elements were found in this cDNA sequence: 147-199, >(GGA
  )n#Simple_repeat
  Seq primer: M13 Forward
  POLYA-No.

FEATURES
  source
    1..260
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:2713294"
      /clone_lib="NCI_CGAP_Sub3"
      /lab_host="DH10B (Life Technologies)"
      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
      polylinker; Site_1: Not I; Site_2: Eco RI; The
      NCI_CGAP_Sub3 library is a subtracted library derived from
      the NCI_CGAP_Sub1 library, which is a subtracted library
      derived from BI. BI constitutes a mixture of 21
      normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4
      , NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
      NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
      NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
      NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Le12,
      NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
      NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
      NCI_CGAP_Brn25. These 21 libraries were pooled and a
      single-stranded DNA preparation of the resulting mixture
      was used as a tracer in a subtractive hybridization with
      a driver whose composition is detailed below:
      NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
      3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775
      , 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342
      , 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831,
      1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1
      LLAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991,
      1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
      3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631,
      1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
      LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids
      985608-986759, 1101192-1101959, 1217928-1220615);

NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
  Clonoids 1057416-1061255, 1144584-1145351). Subtraction
  was performed as previously described [Bonaldo, Lennon &
  Soares (1996)]: Normalization and Subtraction: Two
  Approaches To Facilitate Gene Discovery. Genome Research
  6, 791-806.
TAG_SEQ=None found"
BASE COUNT 30 a 112 c 72 g 46 t
ORIGIN

Query Match      48.4%; Score 15; DB 9; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||
Db 80 GCGCTGAGCAATTG 66

RESULT 12
AI205484/c
LOCUS
DEFINITION
  AI205484
  qr26a12.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:1941982 3'
  similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN
  ); contains element MSR1 MSR1 repetitive element ;, mRNA sequence.
ACCESSION
  AI205484
VERSION
  AI205484.1 GI:3764156
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 292)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished (1997)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
  R. Emmert-Buck, M.D., Ph.D.
  CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 773 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 235.
  Location/Qualifiers
    1..292
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:1941982"
      /clone_lib="NCI_CGAP_GC6"
      /tissue_type="pooled germ cell tumors"
      /lab_host="DH10B"
      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
      polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
      from the normalized library NCI_CGAP_GC4 was prepared, and
      ss circles were made in vitro. Following HAP purification,
      this DNA was used as tracer in a subtractive hybridization
      reaction. The driver was PCR-amplified cDNAs from a pool
      of 5,000 clones made from the same library (clonoids
      1257096-1258631, 1469064-1470983, and 1475592-1476743).
      Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 45 a 123 c 72 g 52 t
ORIGIN
```


Query Match 48.4%; Score 15; DB 9; Length 292;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
| | | | | | | | | | | | | | |
Db 50 GCGCTGAGCAATTG 36

RESULT 13
AA346027/c
LOCUS
DEFINITION
AA346027 Greater omentum I Homo sapiens cDNA 5' end similar to
acetylcholinesterase, mRNA sequence.
ACCESSION
AA346027.1 GI:1998286
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 302)
AUTHORS
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
Gnelm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palannes,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednark,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
,Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
Venter,J.C.
TITLE
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL
Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE
96026280
COMMENT
Other_ESTs: THCl03848
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1. .302
/organism="Homo sapiens"
/db_xref="ATCC (inhost):147585"
/db_xref="taxon:9606"
/clone_lib="Greater omentum I"
/tissue_type="omentum"
/dev_stage="adult"
/note="Organ: NULL; Vector: pBluescript SK-; Site_1: EcoRI
; Site_2: XhoI"
BASE COUNT 60 a 103 c 91 g 44 t 4 others
ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 302;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
| | | | | | | | | | | | | | |
Db 194 GCGCTGAGCAATTG 180

RESULT 14
BE146509/c
LOCUS
DEFINITION
RC1-HT0217-151099-011-g06 HT0217 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BE146509
VERSION
BE146509.1 GI:8609233
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 348)
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-rc1-HT0217-151099-011-g06&t3=1999-10-15&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 348.
Location/Qualifiers
1. .348
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0217"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 122 a 65 c 74 g 87 t
ORIGIN

Query Match 48.4%; Score 15; DB 9; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ggtacctatcaggta 16
| | | | | | | | | | | | | | |
Db 82 GGTACCTATCAGGTA 68

RESULT 15
R28476/c
LOCUS
DEFINITION
R28476 yh50g12.r1 Soares placenta Nb2HP Homo saplens cDNA clone
IMAGE:133222 5', mRNA sequence.

ACCESSION R28476
VERSION R28476.1 GI:784611
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 352)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
, R., Williamson, A., Wohldmann, P. and Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1068
High quality sequence stops: 273
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1068 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 273.
FEATURES
source
1. .352
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:538912"
/db_xref="taxon:9606"
/clone="IMAGE:133222"
/clone_lib="Soares placenta Nb2HP"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AACTGAAGAATTCGCCGCCGAGGAATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 119 a 67 c 88 g 78 t
ORIGIN

Query Match 48.4%; Score 15; DB 10; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggtacctatcaggta 16
|||||
Db 108 GGTACCTATCAGGTA 94

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: August 31, 2002, 22:37:20 ; Search time 154.73 Seconds
(without alignments)
49.212 Million cell updates/sec

Title: US-09-810-861B-2
Perfect score: 31
Sequence: 1 cggtagctatcaggtagcgctgagcaatttg 31

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	51.6	16075	3	US-09-096-942-1
2	16	51.6	16075	3	US-09-096-867-1
3	15	48.4	1845	1	US-07-732-962A-1
4	15	48.4	1845	5	PCT-US92-06106-1
5	15	48.4	2256	2	US-08-318-826A-5
6	15	48.4	2256	2	US-08-370-156-1
7	15	48.4	2256	3	US-08-814-095-1
8	15	48.4	3016	2	US-08-318-826A-7
9	15	48.4	3016	2	US-08-370-156-5
10	15	48.4	3096	2	US-08-814-095-5
11	15	48.4	3096	2	US-08-318-826A-6
12	15	48.4	3096	2	US-08-370-156-3
13	15	48.4	3096	3	US-08-814-095-3
14	15	48.4	35060	3	US-08-814-095-7
15	14	45.2	867	4	US-09-273-839A-7
16	14	45.2	1920	1	US-08-746-789A-1
17	14	45.2	6048	4	US-09-634-920-3
18	14	45.2	17138	4	US-09-813-819-3
19	14	45.2	17138	4	US-09-920-048-3
20	13	41.9	289	2	US-08-967-101-23
21	13	41.9	289	2	US-08-592-541-23
22	13	41.9	289	3	US-09-124-698-23
23	13	41.9	289	4	US-09-127-480-23
24	13	41.9	289	4	US-08-496-841C-23
25	13	41.9	335	2	US-08-623-906A-12
26	13	41.9	401	3	US-08-581-918A-27
27	13	41.9	401	4	US-08-346-147B-27

28	13	41.9	404	3	US-08-581-918A-26	Sequence 26, Appl
29	13	41.9	404	4	US-08-346-147B-26	Sequence 26, Appl
30	13	41.9	837	4	US-09-570-367C-20	Sequence 20, Appl
31	13	41.9	837	4	US-09-570-367C-22	Sequence 22, Appl
32	13	41.9	1149	1	US-08-474-177-3	Sequence 3, Appl
33	13	41.9	1149	1	US-08-487-033-3	Sequence 3, Appl
34	13	41.9	1149	1	US-08-480-810-3	Sequence 3, Appl
35	13	41.9	1149	2	US-08-508-735-3	Sequence 3, Appl
36	13	41.9	1149	2	US-08-848-251-3	Sequence 3, Appl
37	13	41.9	1149	2	US-08-486-047-3	Sequence 3, Appl
38	13	41.9	1149	3	US-09-120-130-3	Sequence 3, Appl
39	13	41.9	1149	3	US-09-115-252-3	Sequence 3, Appl
40	13	41.9	1149	3	US-08-986-515-3	Sequence 3, Appl
41	13	41.9	1149	3	US-09-120-128-3	Sequence 3, Appl
42	13	41.9	1149	4	US-09-120-129-3	Sequence 3, Appl
43	13	41.9	1149	4	US-09-201-139-3	Sequence 3, Appl
44	13	41.9	1149	4	US-09-120-131-3	Sequence 3, Appl
45	13	41.9	1608	1	US-08-393-985-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-096-942-1
; Sequence 1, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,942
; EARLIER FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16075
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; PUBLICATION INFORMATION:
US-09-096-942-1

Query Match 51.6%; Score 16; DB 3; Length 16075;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 tcaggtagcgctgagc 25
|||||
Db 11351 tcaggtagcgctgagc 11366

RESULT 2
US-09-096-867-1
; Sequence 1, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,867

; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 16075
; TYPE: DNA
; ORGANISM: Xanthomonas campestris
; PUBLICATION INFORMATION:
US-09-096-867-1

Query Match 51.6%; Score 16; DB 3; Length 16075;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 tcaggtagcgcgtgagc 25
|||||

Db 11351 tcaggtagcgcgtgagc 11366

RESULT 3
US-07-732-962A-1/c
; Sequence 1, Application US/07732962A
; Patent No. 5248604
; GENERAL INFORMATION:
; APPLICANT: Fischer, Meir
; TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07732,962A
; FILING DATE: 19910722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842

US-07-732-962A-1

Query Match 48.4%; Score 15; DB 1; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31

Db 1718 GCGCTGAGCAATTG 1704
|||||

RESULT 4
PCT-US92-06106-1/c
; Sequence 1, Application PC/TUS9206106
; GENERAL INFORMATION:
; APPLICANT: Fischer, Meir
; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06106
; FILING DATE: 19920722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842

PCT-US92-06106-1

Query Match 48.4%; Score 15; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||

Db 1718 GCGCTGAGCAATTG 1704

RESULT 5
US-08-318-826A-5/c
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410

Query Match 48.4%; Score 15; DB 1; Length 1845;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31

;; CITY: Farmington Hills
;; STATE: Michigan
;; COUNTRY: US
;; ZIP: 48334
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/318,826A
;; FILING DATE:
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: 2391.00001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 539-5050
;; TELEFAX: (248) 539-5055
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2256 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA to mRNA
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
;; OTHER INFORMATION: 3, 4 and 6"
US-08-318-826A-5

Query Match 48.4%; Score 15; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 6
US-08-370-156-1/c
; Sequence 1, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (810) 689-3500
;; TELEFAX: (810) 689-4071
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2256 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-370-156-1

Query Match 48.4%; Score 15; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 7
US-08-814-095-1/c
; Sequence 1, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "ACHE gene comprising exons
; DESCRIPTION: 2, 3, 4 and 6"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-814-095-1

```
Query Match      48.4%; Score 15; DB 3; Length 2256;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 17 gcgctgagcaatttg 31
 |||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 8
 US-08-318-826A-7/C
 ; Sequence 7, Application US/08318826A
 ; Patent No. 5891725
 ; GENERAL INFORMATION:
 ; APPLICANT: Soreq, Hermona
 ; APPLICANT: zakut, Haim
 ; APPLICANT: Eckstein, Fritz
 ; TITLE OF INVENTION: Synthetic Antisense
 ; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
 ; TITLE OF INVENTION: Containing Them
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kohn & Associates
 ; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
 ; CITY: Farmington Hills
 ; STATE: Michigan
 ; COUNTRY: US
 ; ZIP: 48334
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/318,826A
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kohn, Kenneth I.
 ; REGISTRATION NUMBER: 30,955
 ; REFERENCE/DOCKET NUMBER: 2391.00001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (248) 539-5050
 ; TELEFAX: (248) 539-5055
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3016 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 160..2010
 ; OTHER INFORMATION: /note= "Splice Variant: Exons 1, 2,
 ; OTHER INFORMATION: 3, 4, 5 and 6"
 ;
 ; US-08-318-826A-7

```
Query Match      48.4%; Score 15; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 17 gcgctgagcaatttg 31
 ++++++
Db 1877 GCGCTGAGCAATTG 1863

RESULT 9
US-08-370-156-5/c
; Sequence 5, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
US-08-370-156-5

Query Match	48.4%	Score 15;	DB 2;	Length 3016;
Best Local Similarity	100.0%	Pred. No. 2.6;		
Matches 15;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 17 gcgctgagcaatttg 31
 |||
 Db 1877 GCGCTGAGCAATTG 1863

RESULT 10
US-08-814-095-5/c
; Sequence 5, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.

; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced Ache
; DESCRIPTION: comprising exons 2, 3, 4, 5 and 6"
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; US-08-814-095-5

Query Match 48.4%; Score 15; DB 3; Length 3016;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 11
US-08-318-826A-6/c
; Sequence 6, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.

; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4, 5 and the translated portion of Intron 4 (readthrough
; US-08-318-826A-6

Query Match 48.4%; Score 15; DB 2; Length 3096;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 12
US-08-370-156-3/c
; Sequence 3, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
US-08-370-156-3

Query Match      48.4%; Score 15; DB 2; Length 3096;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||
Db 1877 GCGCTGAGCAATTG 1863

RESULT 13
US-08-814-095-3/c
; Sequence 3, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced AChE
; DESCRIPTION: comprising exons 2, 3, 4 and 5 as well as the translated portion
; DESCRIPTION: of Intron 4 (readthrough)"
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
US-08-814-095-3

Query Match      48.4%; Score 15; DB 3; Length 3096;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gcgctgagcaatttg 31
|||||

```

```

Db 1877 GCGCTGAGCAATTG 1863

RESULT 14
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Cosmid including ACHE
; DESCRIPTION: promoter, ACHE gene and ARS gene"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7q22
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 24090..25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "(translation start:
; OTHER INFORMATION: 24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:

```

```

; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 27385..27387
; FEATURE:
; NAME/KEY: exon
; LOCATION: 28008..28129
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 28129..28131
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34528..34895)
; OTHER INFORMATION: /function= "arsenite resistance"
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34092..34358)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33779..33963)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33493..33591)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33297..33408)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32959..33094)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32569..32628)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32386..32468)

```

```

; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31894..32080)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 9
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31363..31534)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31131..31284)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "ARS"
; OTHER INFORMATION: /number= 16
; US-08-814-095-7
;
; Query Match 48.4%; Score 15; DB 3; Length 35060;
; Best Local Similarity 100.0%; Pred. No. 2.5;
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 17 gcgctgagcaatttg 31
; |||||
; Db 27168 GCGCTGAGCAATTG 27154
;
; RESULT 15
; US-09-273-839A-7
; Sequence 7, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Bacillus anthracis

```

US-09-273-839A-7

Query Match 45.2%; Score 14; DB 4; Length 867;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 tagcgctgagcaat 28
|||||
Db 852 tagcgctgagcaat 865

Search completed: August 31, 2002, 22:37:24
Job time: 19308 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:35:31 ; Search time 9324.54 Seconds
(without alignments)
12942.551 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcattgcctgcaggtc.....ctatgaccatgattacgccca 5767

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description

1	5767	100.0	5767	6	AX275254	AX275254 Sequence
2	2637	45.7	3858	6	AX114873	AX114873 Sequence
3	2637	45.7	4473	12	UCU89964	U89964 Cloning vec
4	2637	45.7	4960	6	AX191663	AX191663 Sequence
5	2637	45.7	5711	6	AX114861	AX114861 Sequence
6	2637	45.7	7332	6	AX191664	AX191664 Sequence
7	2637	45.7	8062	6	AX114872	AX114872 Sequence
8	2637	45.7	8153	6	AX114871	AX114871 Sequence
9	2635	45.7	4518	12	XXP35SGFP	U28417 Cloning vec
10	2635	45.7	4519	12	U02456	U02456 Cloning vec
11	2635	45.7	4674	12	XXU02437	U02437 Cloning vec
12	2630	45.6	3796	6	E49323	E49323 Infectious
13	2630	45.6	5452	12	U02454	U02454 Cloning vec
14	2630	45.6	6541	12	XXU02433	U02433 Cloning vec
15	2630	45.6	6894	12	XXU02435	U02435 Cloning vec
16	2630	45.6	7069	12	U02442	U02442 Cloning vec
17	2630	45.6	7164	6	AX195206	AX195206 Sequence
18	2630	45.6	7164	12	U02451	U02451 Cloning vec
19	2630	45.6	7481	12	XXU02438	U02438 Cloning vec
20	2628	45.6	4311	6	AX244155	AX244155 Sequence
21	2628	45.6	4950	6	AR119454	AR119454 Sequence
22	2628	45.6	5733	8	ANGPDAG	Z32524 A.nidulans
23	2628	45.6	5760	12	EVPAN24	Z32750 Expression
24	2628	45.6	9780	6	A82653	A82653 Sequence 3
25	2628	45.6	11099	6	E16636	E16636 Expression
26	2628	45.6	11099	6	E16677	E16677 All sequenc
27	2628	45.6	12494	8	D84238	D84238 Nicotiana t
28	2628	45.6	15528	6	A93016	A93016 Sequence 4
29	2628	45.6	15528	12	PEAVGEN	Y07862 Cloning vec
30	2627	45.6	5670	12	EVPAN522	Z32688 Expression
31	2627	45.6	5670	12	EVPAN523	Z32689 Expression
32	2627	45.6	5721	12	EVPAN521	Z32697 Expression
33	2627	45.6	5748	12	EVPAN534	Z32699 Expression
34	2627	45.6	6089	12	EVPAN81	Z32751 Expression
35	2627	45.6	6756	12	EVPAN71	Z32698 Expression
36	2627	45.6	7599	12	EVPNO102	Z32701 Expression
37	2621	45.4	7596	12	AF347016	AF347016 Shuttle v
38	2618	45.4	2854	12	CVU34922	U34922 Cloning vec
39	2613	45.3	3024	1	PK8CCDB	L27083 Plasmid KIL
40	2613	45.3	3024	12	SYNCCDB	L38496 Cloning vec
41	2613	45.3	4342	6	AX244154	AX244154 Sequence
42	2586	44.8	4392	12	CVU47294	U47294 Cloning vec
43	2585	44.8	5027	12	AF276982	AF276982 Integrati
44	2584	44.8	2743	12	CVPGEM3Z	X65304 Cloning vec
45	2584	44.8	2746	12	CVPGEM4Z	X65305 Cloning vec

ALIGNMENTS

RESULT	1
AX275254	AX275254
LOCUS	Sequence 3 from Patent WO0171014.
DEFINITION	AX275254
ACCESSION	AX275254
VERSION	AX275254.1 GI:16547674
KEYWORDS	synthetic construct.
SOURCE	synthetic construct
ORGANISM	artificial sequence.
REFERENCE	1 (sites)
AUTHORS	Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE	Expression of recombinant human acetylcholinesterase in transgenic plants
JOURNAL	Patent: WO 0171014-A 3 27-SEP-2001; BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor, Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason, Hugh S. (US)
FEATURES	Location/Qualifiers
source	1. .5767 /organism="synthetic construct" /db_xref="taxon:32630" /note="plasmid vector pTM034"

BASE COUNT		1390 a	1521 c	1495 g	1361 t		
ORIGIN							
Query Match		100.0%;		Score 5767;		DB 6; Length 5767;	
Best Local Similarity		100.0%;		Pred. No. 0;			
Matches 5767;		Conservative 0;		Mismatches 0;		Indels 0; Gaps 0;	
QY	1	agcttgcatgcctgcaggtcaacatggtggagcagcacacactctcgtctactccaagaata	60				
Db	1	AGCTTGCATGCCCTGCAGGTCAACATGGTGGAGCAGCACACTCTCGTCTACTCCAAGAATA	60				
QY	61	tcaagatacagtgctcagaagaccagagggctattgagacttttcaacaagggtaatat	120				
Db	61	TCAAAGATACAGTCTCAGAAGACCAGAGGGCTATTGAGACTTTTCAACAAAGGGTAATAT	120				
QY	121	cgggaaacctcctcggaattccattgccccagctatctgtcacttcacgcgaagacagtag	180				
Db	121	CGGGAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTCATCGAAAGGACAGTAG	180				
QY	181	aaaaggaagatggcttctacaaaatgccatcattgcgataaaaggaaggctatcgttcaag	240				
Db	181	AAAAGGAAGATGGCTTCTACAAATGCCATCATTTGGGATAAAGGAAAGGCTATCGTTCAAG	240				
QY	241	aatgcctctaccgacagtggtcccaaaagatggagccccccacccacgagggaacatcgtggaa	300				
Db	241	AATGCCTCTACCGACAGTGGTCCCAAAGATGGACCCCCCACCACGAGGAACATCGTGGAA	300				
QY	301	aaagaagacgttccaaccacgtcttcaaaagcagtggaattgatgtataaacttttcaaca	360				
Db	301	AAAGAAGACGTTCCAACCACGCTCTTCAAAGCAAGTGGATTGATGTGATAACTTTTCAACA	360				
QY	361	aagggtaatatcgsgaaacctcctcggattccattgcccagctatctgtcacttcatcga	420				
Db	361	AAGGGTAATATCGGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTCATCGA	420				
QY	421	aaggacagttagaaaaggaagatggcttctacaaaatgccatcattgcgataaaaggaaggc	480				
Db	421	AAGGACAGTAGAAAGGAAGATGGCTTCTACAAATGCCATCATTTGCCGATAAAGGAAGGC	480				
QY	481	tatcgttcaagaatgcctctaccgacagtggtcccaaaagatggacccccccacccacgagga	540				
Db	481	TATCGTTCAAGAATGCCTCTACCGACAGTGGTCCCAAAGATGGACCCCCCACCACGAGGA	540				
QY	541	acatcgtggaaaaaagagacgttcccaaccacgtcttcaaaagcaagtggattgatgtgata	600				
Db	541	ACATCGTGGAAAAAAGAGACGTTCCAACCACGCTTTCAAAGCAAGTGGATTGATGTGATA	600				
QY	601	tctccactgacgtaggggatgacgcacaaatcccactatccttcgcaagacccttcctcta	660				
Db	601	TCTCCACTGACGTAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTA	660				
QY	661	tataaggaagtctattcatttggagaggacctcgagaattaattctcaacacaacatat	720				
Db	661	TATAAGGAAGTTCATTTCATTGGAGAGGACCTCGAGAATTAAATCTCAACACAACATAT	720				
QY	721	acaaaaacaaatctcaagcaatcaagcattctacttctatttgacgaatttaaatca	780				
Db	721	ACAAAACAACGAATCTCAAGCAATCAAGCATTTCTACTTCTATTGCAGCAATTTAAATCA	780				
QY	781	tttcttttaaaagcaaaagcaattttctgaaaaattttcaccatttacgaaacgatagccatg	840				
Db	781	TTTTCTTTTAAAGCAAAAGCAATTTTCTGAAAATTTTCACCATTTACGAACGATAGCCATG	840				
QY	841	gctcccccgagtgctgctgcacacgccttcctcgcttccccactccttctcctcctc	900				
Db	841	GCTCCCCCGCAGTGTGCTGCACACGCGCTTCCTGGCTTCCCCACTCCTTCTCCTCCTC	900				
QY	901	ctctggctcctgggtggagagatggggcgtagggccgggaggatgcagagctgctgggtg	960				
Db	901	CTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTGGTG	960				
QY	961	acggtgcgtggggcgcgctgcggggcattcgcctgaagacccccgggggcctgtctct	1020				

Db	961	ACGGTGCCTGGGGCGCGGCTGCGGGCATTCGCTGAAGACCCCGGGGCGCCTGTCTCT	1020
QY	1021	gctttcctgggcatcccccttgcggagccaccatgggacccccgcctcttctgcccaccg	1080
Db	1021	GCTTTCCTGGGCATCCCTTTGCGGAGCCACCCTAGGACCCCGTCGCTTCTGCCACCG	1080
QY	1081	gagcccaagcagccttggtcaggggtggtagacgctacaaccttccagagtgtctgctac	1140
Db	1081	GAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGTAC	1140
QY	1141	caatatgtggacacccctatacccagggttttgaggggcaccgagatgtggaacccccaccgt	1200
Db	1141	CAATATGTGGACACCCCTATACCCAGGTTTTTGAGGGCACCCGAGATGTGGAACCCCAACCGT	1200
QY	1201	gagctgagcagaggactgcctgtacctcaacgtgtggacaccatacccccgccctacatcc	1260
Db	1201	GAGCTGAGCGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCGGCTACATCC	1260
QY	1261	cccacccctgtcctcgtctggtgatctatgggggtggtcttctacagtggggcctcctccttg	1320
Db	1261	CCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCCTTG	1320
QY	1321	gacgtgtacgatggcgccttcttggtacagggccgagaggaactgtgctggtgtccatgaac	1380
Db	1321	GACGTGTACGATGGCGCTTCTTGGTACAGGCCCGAGAGGACTGTGCTGGTGTCCATGAAC	1380
QY	1381	taccgggtgggagcctttggcttccctgcctgcctgcctgcctgcctgcctgcctgcctgcct	1440
Db	1381	TACCGGGTGGGAGCCTTTGGCTTCTTGGCCCTGCCGGGAGCCGAGAGGCCCGGGGCAAT	1440
QY	1441	gtgggtctcctcgtgatacagaggtggccctgcctgcctgcctgcctgcctgcctgcctgcct	1500
Db	1441	GTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCCTTC	1500
QY	1501	gggggtgaccccgacatcagtgacgtgttggggagagcgcgggagcgcctcgtggggc	1560
Db	1501	GGGGGTGACCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGAGCCGCTCGGTGGGC	1560
QY	1561	atgcacctgctgtcccgcgccacgcgggcctgttccacagggcctgctgcagagcggt	1620
Db	1561	ATGCACCTGCTGCCCGCCACGCGGGGCCCTGTTCACAGGGCCGTGCTGCAGAGCGGT	1620
QY	1621	gcccccaatggacctgggccacggtgggcatgggagagggccctgcgagggccacgcag	1680
Db	1621	GCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCGTCGCGAGGCCACGCAG	1680
QY	1681	ctggccccacctgtgggctgtcctccaggcggcactggtgggaatgacacagagctggtta	1740
Db	1681	CTGGCCCCACCTTGTGGGCTGTCTCTCCAGGGCGGCACTGGTGGGAATGACACAGAGCTGGTA	1740
QY	1741	gcctgccttcggacacgaccagcgcaggctcctcgtgcctgtggtagaatggcagctgcct	1800
Db	1741	GCCTGCCTTCGGACACGACCCAGCGCAGGTCTCTGGTGAACCAACGAAATGGCACGTGCTGCCT	1800
QY	1801	caagaaagcgtcttcgggttctcctcgtgcctgtggtagaatggagacttccctcagtgac	1860
Db	1801	CAAGAAAGCGTCTCCGGTCTCTCTTCGTGCTGTGGTAGATGGAGACTTCCTCAGTGAC	1860
QY	1861	acccagaggccctcatcaacgcgggagacttccacggccctgcaggtgtggtgggtgtg	1920
Db	1861	ACCCAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGTGTG	1920
QY	1921	gtgaaggatgagggtcgtattttctggtttacgggggccccaggcttcagcaaaagacaac	1980
Db	1921	GTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAAAGACAAC	1980
QY	1981	gagtctctcatcagccgggcccagagttcctgcgcgggggtgcgggtgtccccagggtta	2040
Db	1981	GAGTCTCTCATCAGCCGGGCCGAGTTCTTGCCGGGGTGCGGGTCCGCCAGGTA	2040
QY	2041	agtgacctggcagccgaggtgtggtcctgcattacacagactggtgcatacccgaggac	2100

Db 2041 AGTGACCTGGCAGCCGAGGCTGTGTCCTGTCAATTACACAGACTGGCTGCATCCCGAGGAC 2100

Qy 2101 ccggcagccctgagggagggccctgagcgatgtggtgggcgaccacaatatgcgtgtgcccc 2160

Db 2101 CCGGCACGCCTGAGGGAGGCCCCTGAGCGATGTGGTGGCGACCAACAATGTCGTGCCCC 2160

Qy 2161 gtggccagctggctgggcgactggctgcccagggtgcccgggtctacgcctacgtcttt 2220

Db 2161 GTGGCCAGCTGGCTGGCGACTGGTGGCCAGGGTGCCCGGGTCTACGCCCTACGTCCTT 2220

Qy 2221 gaacacgctgtctccacgctctcctggcccccctgtggaatgggggtgccccacggctacgag 2280

Db 2221 GAACACCGTGCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGTGCCCCACGGCTACGAG 2280

Qy 2281 atcgagttcatctttgggataccccctggaccccctctcgaaaactacacggcagagagaaa 2340

Db 2281 ATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAGAAA 2340

Qy 2341 atcttgcgccagcgactgatgcgatactgggccaaactttgcccgcacaggggatcccaat 2400

Db 2341 ATCTTCGCCCAGCGACTGATGCGATACTGGGCCAACTTTGCCCCACAGGGGATCCCAAT 2400

Qy 2401 gagcccgagagccccaagggccccaacaatggccccctacacgggggggctcagcagttac 2460

Db 2401 GAGCCCGGAGACCCCAAGGCCCCACAAATGGCCCCCGTACACGGSGGGGCTCAGCAGTAC 2460

Qy 2461 gttagctcggacctgcggccgctggagggtgcggcggggctgcgcgccccagccctgcgcc 2520

Db 2461 GTTAGTCTGGACCTGGCGCGCTGGAGGTGGCGGGGCTGGCGGCCACAGCCCTGGGCC 2520

Qy 2521 ttctggaaccgcttcctcccccataatgctcagcgctacacctgataggtaccgagctctctc 2580

Db 2521 TTCTGGAAACCGCTTCTTCCCCAAATTGCTCAGCGCTACCTGATAGGTACCGAGCTCTCTC 2580

Qy 2581 aacaatctagctagagtttgcctccatactatatgtaataaggtgatgctgatgcactat 2640

Db 2581 AACAACTAGCTAGAGTTTGCTCCCTATCTATATGTAATAAGGTATGCTGATATGCACTAT 2640

Qy 2641 tcaatagggagcattagctatgttgtaagtctcactttatgttatgtgggtaagtcaac 2700

Db 2641 TCAAATAGGAGCATTAGCTATGTTGTTAATGTCACTTTATGTTATGTGGGTAAGTCACC 2700

Qy 2701 taagcacctccacgtacctaactgttgctctcttaccggccttaataaatctctcgccct 2760

Db 2701 TAAGACACTCCACGTACCTACGTTGTGTCTCTTACCGGCTTTAATAAATCTTCTGCCCT 2760

Qy 2761 tgttccatattactaataattatccctttcttcaactaaaaaagaaattggttatcattaaagtat 2820

Db 2761 TGTTCCATATTTACTAATTATCCCTTCTTCACCTAAAAGAAATTTGTTATCATTTAAGTAT 2820

Qy 2821 tagtctttagaacacatatgaggtctttaattgggtaggtttttacaaaattaaataataaaa 2880

Db 2821 TAGTCTTTAGAACACATATGAGGCTCTTTAATTGGGTAGGTTTTTACAAATTAACATAATAA 2880

Qy 2881 atgtcataaaaatccacgtgtttaacaaaatgcagaaaatcgacgtcgtcttatggaccga 2940

Db 2881 ATGTCATAAAATCCACGTGGTTAAACAAATGCAGAAATCGACGTCGTCTATTGGACCGA 2940

Qy 2941 cagttgctattataataatgggccaccatagtagactgacaaaataaattaccctgacaaaca 3000

Db 2941 CAGTTGCTATTATAATATGGGCCACCATAGTAGACTGACAAATAAATTACCTGACAACA 3000

Qy 3001 tcggttccactaaaataacaaaacacaaaaggagtgcatctttccaggggcatcttttgtaata 3060

Db 3001 TCGTTTCACTAATAACAAACACAAAAGGGAGTGCAATTTTCCAGGGCATTTTGTGTAATA 3060

Qy 3061 aaaaaaggttaaaaggagtgcaatagaaaaatataggggtgtggaatatagtattgagca 3120

Db 3061 AAAAACAGTTAAAGGGAGTGCATAGAAATATAGGGGTGTGGAAATAGTGATTGTAGCA 3120

Qy 3121 cgtcttgagcgcaattcactggccgtcgcttttacaacgtcgtgactgggaaaaacccctggc 3180

Db 3121 CGTCTTGAAGCGAATTCACTGGCCGTCGTTTACAACGTCGTGACTGGGAAACCCCTGGC 3180

Qy 3181 gttacccaacttaatcgcccttgcgagcacatccccctttcgccagctggcgtaaatagcgaa 3240

Db 3181 GTTACCCAACCTAATCGCCTTGCGAGCACATCCCCCTTTTCGCCAGCTGCGGTAATAGCGAA 3240

Qy 3241 gagggccgcacgatcgcccttcccaacagttgcgcagccctgaatggcgaatggcgccctg 3300

Db 3241 GAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTG 3300

Qy 3301 atgcgggtattttctccttacgcatactgtgcgggtattttcacacccgcatatgggtgcactctc 3360

Db 3301 ATGCGGTATTTTCTCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACTCTC 3360

Qy 3361 agtacaatctgctctctgatgcgcgcatagttaagccagccccgcagaccccgccaacacccgct 3420

Db 3361 AGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCCGCAACACCCGCT 3420

Qy 3421 gaocggccctgaagggttctctgctccggcatccgcttacagacagctgtgaccgctc 3480

Db 3421 GACGGCCCTGACGGGCTTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTC 3480

Qy 3481 tccgggagctgcatgtgcagaggttttcccgctcatcacccgaaacgcgcgagacgaaaag 3540

Db 3481 TCCGGGAGCTGCATGTGCAGAGGTTTACCGCTCATCACCGAAACGCGGAGACGAAAG 3540

Qy 3541 ggcctcgtgatacgcctatttttataggttaatgcatgataataatggtttcttagacg 3600

Db 3541 GGCCCTCGTGATACGCCCTATTTTATAGGTTAATGTCTATGATAATAATGTTTCTTAGACG 3600

Qy 3601 tcagggtggcacttttcggggaatgtgcggggaacccctatttgtttatttttctaaata 3660

Db 3601 TCAGGTGGCACTTTTCGGGGAAATGTGCGGGAAACCCCTATTTGTTATTTTCTAAATA 3660

Qy 3661 cattcaaaatagtatccgctcatgagacaataaacctgataaaatgcttcaataatattga 3720

Db 3661 CATTCAAATATGTATCCGCTCATGAGACAATAAACCTGATAAAATGCTTCAATAATATTGA 3720

Qy 3721 aaaagggaagtagtagtattcaacatttcggtgcgccttatcccttttttgcggca 3780

Db 3721 AAAAGGAAGAGTATGAGTATTCAACATTTCCGCTGTCGCCCTTATTCCCTTTTTCGGGCA 3780

Qy 3781 tttgcctcctgttttgcacccagaaacgctggtgaaagttaaagatgctgaagat 3840

Db 3781 TTTTGGCTTCCGTGTTTGTCTACCCAGAACGCTGGTGAAGTAAAAGATGCTGAAGAT 3840

Qy 3841 cagttgggtgcacgagtggtttacatcgaactggatctcaacagcggtaagatcccttgag 3900

Db 3841 CAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAG 3900

Qy 3901 agttttcgccccgaagaacgttttccaatgatgagcaccttttaaagttctgctatgtggc 3960

Db 3901 AGTTTTCGCCCGGAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGC 3960

Qy 3961 gcgggtattatcccgtaattgacgcggggcagaagcaactcgtgcgcgacatacactattct 4020

Db 3961 GCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCCGCCCATACACTATTCT 4020

Qy 4021 cagaatgacttggttagtactcaccagtcacagaaaaagcatcttacggtggcatgaca 4080

Db 4021 CAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAAGCATCTTACGGATGGCATGACA 4080

Qy 4081 gtaagagaatttatgcagtgctgccataaacatgagtataacacactgcggccaacttactt 4140

Db 4081 GTAAGAGAAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTCGGGCCAACCTACTT 4140

Qy 4141 ctgacaacgatcggaggaccgaaggagctaacccgcttttttgcaacaacatgggggatcat 4200

Db 4141 CTGACAACGATCGGAGGACCGGAAGGAGCTAACCCGCTTTTTCACACAACATGGGGGATCAT 4200

Qy 4201 gtaactcgccctgatcgttgggaacccggagctgaatgaagccataaccaacgacgagcgt 4260

Db 4201 GTAACTCGCCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGT 4260

QY 4261 gacacacgatgccttagcaaatggcaacaacgttgcgcaaaactattaaactggcgaaacta 4320
Db 4261 GACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGCGCAACTA 4320
QY 4321 cttactctagcttcccgcaacaattaatagactggatggaggcgataaaagtgcagga 4380
Db 4321 CTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGA 4380
QY 4381 ccacttctgcgctcgcccttccggctgctggtttattgctgataaaatctggagccggt 4440
Db 4381 CCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGGT 4440
QY 4441 gagcgtgggtctcgcggtatcatctgagcacctggggccagatggtaagccctcccgatc 4500
Db 4441 GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGSCCAGATGGTAAGCCCTCCCGTATC 4500
QY 4501 gtagttatctacacgaggggagtcaggcaactatggatgaacgaatagacagatcgct 4560
Db 4501 GTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCT 4560
QY 4561 gagatagtgccctcactgattaaagcattggtaactgtcagaccaagtcttactcatata 4620
Db 4561 GAGATAGTGCCCTCACTGATTAAAGCATTTGGTAAGTGTCTCAGACCAAGTTTACTCATATA 4620
QY 4621 ctttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttt 4680
Db 4621 CTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 4680
QY 4681 gataatctcatgacaaaaatcccttaacgtgagtttctgctccactgagcgcgaacccc 4740
Db 4681 GATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCTACTGAGCGTCAGACCCC 4740
QY 4741 gtgaaaaagatcaaagatctcttgagatccctttttctgcgctaactgctgcttg 4800
Db 4741 GTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTCTGCGGTAATCTGCTGTG 4800
QY 4801 caaacaacaaacccacgcgtaccagcggtgttctggttcgggatcaagagctaccaact 4860
Db 4801 CAAACAACAAACCCACCGCTACACGGTGTGTTTGTTCGGGATCAAGAGTACCAACT 4860
QY 4861 ctttttccgaaggtaactgcttcagcagagcgagataccaataactgtcctcttagtg 4920
Db 4861 CTTTTCGGAAGGTAAGTGGCTTCAGCAGAGCGCGAGATACCAATACTGTCTTCTAGTG 4920
QY 4921 tagccgtagttaggccaccacttcaagaactctgtagcacccgctacataactcgctctg 4980
Db 4921 TAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCTG 4980
QY 4981 ctaatcctgttaccagtggctgctgccagtgccagtgagataagtcgtgtcttacgggtggac 5040
Db 4981 CTAATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGTTGGAC 5040
QY 5041 tcaagacgatagttacgggataaaggcgagcggtcggtgaacgggggttcgtgcaca 5100
Db 5041 TCAAGACGATAGTTACCGGATAAGGCGCAGCGTCGGGCTGAACGGGGGTTTCGTGCACA 5100
QY 5101 cagcccagcttgagggaacgacctacacccaactgagataacctacagcgtgagctatga 5160
Db 5101 CAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGA 5160
QY 5161 gaaagcggccacgcttcccgaaaggagaaaggcgagcaggtatccgggtaagcgggcggtc 5220
Db 5161 GAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAAGCGGACGGTC 5220
QY 5221 ggaacaggagagcgacagaggagcttccaggggaaacgcctggtatctttatagtcct 5280
Db 5221 GGAACAGGAGAGCGCAGGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCCT 5280
QY 5281 gtcgggtttccgccactctgacttgagcgtcgatttttctgatgctcgtcagggggcg 5340
Db 5281 GTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGG 5340
QY 5341 agcctatggaaaaacgccagcaacggcgccctttttacgggttcctggccttttctgctgcct 5400

Db 5341 AGCCTATGGAACAAACGCGAGCAACGCGCCTTTTACGGTTCTTGGCCTTTTGGTGCCT 5400
QY 5401 tttgctcacatgttcttctgctgttatccccctgattctgtggataaacctgattaccgccc 5460
Db 5401 TTTGCTCACATGTTCTTCTGCGTTATCCCTGTATCTGTGGATAACCGTATTACCGCC 5460
QY 5461 tttgagtgagctgatacgcgtcgccgcagccgaaacgagcagcagcagcagcagcagc 5520
Db 5461 TTTGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGTCAGTGAGC 5520
QY 5521 gaggaagcggaagagcgcccaatacgcgaacacgcctctccccgcgcgttgccgagattcat 5580
Db 5521 GAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCTTGGCCGATTCAT 5580
QY 5581 taatgcagctggcacagcaggtttcccgactggaaagcgggcagtgagcgcgaacgcaatt 5640
Db 5581 TAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATT 5640
QY 5641 aatgtgagttagctcactcattagcaccacccaggctttacactttatgcttcggctcgt 5700
Db 5641 AATGTGAGTTAGTCACTCATTAGGCACCCAGGCTTTACACTTTTATGCTTCCGCTCGT 5700
QY 5701 atgtgtgtggaattgtgagcggataaacaatttcacacaggaacagctatgaccatgat 5760
Db 5701 ATGTTGTGTGAATTGTGAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGAT 5760
QY 5761 tacgcca 5767
Db 5761 TACGCCA 5767

RESULT 2
AX114873
LOCUS AX114873 Sequence 33 from Patent WO0129208. DNA linear PAT 11-MAY-2001
DEFINITION AX114873
ACCESSION AX114873
VERSION AX114873.1 GI:14031815
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 3858)
AUTHORS Kuehn, R., von Melchner, H. and Altschmied, J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 33 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE); Frankgen Biotechnologie AG (DE)
FEATURES
source
1. .3858
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector PRK50"

BASE COUNT 943 a 959 c 994 g 962 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactggccgctgcttttacaaacgtcgtgactgggaaaaacccctggcggttaccacaac 3190
Db 1187 CGAATTCACTGGCCGTCGTTTACAAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAAC 1246
QY 3191 ttaatcgcccttgacgacacatccccctttcgccagctggcgtaatagcgaagagggcccgca 3250
Db 1247 TTAATCGCCTTGACGACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 1306
QY 3251 ccgatcgcccttcccaacagttgcgcagcctgaatggcggaatggcgccctgatgcggtatt 3310
Db 1307 CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATT 1366
QY 3311 ttctccttacgcattctgtgcggtattttcacacgcgcataatggtgcactctcagtcacaattct 3370

2638/38

Db	1367	TTCCTCTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCT	1426
Qy	3371	gctctgagtcgcgcatagtttaagccagccccgacacacccgcgaacaccccgctgacgcgcct	3430
Db	1427	GCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCT	1486
Qy	3431	gacgggcttgctgctcccgccatccgcgttacagacaagctgtgaccgtctccggagct	3490
Db	1487	GACGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTACCCGTCTCCGGGAGCT	1546
Qy	3491	gcattgtgcagaggttttcaccgctcatcaccgaacgcgcgagacgaagggcctcgtga	3550
Db	1547	GCATGTGTCAAGGTTTTTACCCTCATCACCGAAACGCGCGAGACGAAGGGCCTCGTGA	1606
Qy	3551	tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggca	3610
Db	1607	TACGCCATATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAAGTGGCA	1666
Qy	3611	cttttcgggggaaatgtgcgcgggaacccctatttgtttattttttctaaatacatattcaaaata	3670
Db	1667	CTTTTTCGGGGAAATGTGCGCGGAACCCCTATTGTGTTATTTTCTAAATACATTCAAAATA	1726
Qy	3671	tgtatccgctcatgagacaaataaaccttgataaaatgcttcaataataattgaaaaagggaaga	3730
Db	1727	TGTATCCGCTCATGAGACAAATAACCTTGATAAAATGCTTCAATAATATTGAAAAAGGAAGA	1786
Qy	3731	gtatgagtattcaacatttcctgctgcgccttattcccttttttgcggcattttgccttc	3790
Db	1787	GTATGAGTATTCAACATTTCCGTGTCCCTTATTCCCTTTTTCGGGCATTTTGCCTTC	1846
Qy	3791	ctgttttttgctcaccacagaaacgcgtggtgaagtaaaagatgctgaagatcagttgggtg	3850
Db	1847	CTGTTTTTGTCTCACCCAGAAACGCTGCTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG	1906
Qy	3851	cacgagtgggttacatcgaaactggatctcaacagcggtaagatcccttgagagttttcgcc	3910
Db	1907	CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC	1966
Qy	3911	ccgaagaacgttttccaaatgatgagcaacttttaaagttctgctatgtggcgcggtattat	3970
Db	1967	CCGAAGAAGGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTATTAT	2026
Qy	3971	cccgtattgacccgggcaagagcaactcggtcgcgcgcatatacctattctcagaatgact	4030
Db	2027	CCCGTATTGACGCCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACT	2086
Qy	4031	tggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaat	4090
Db	2087	TGGTTGAGTACTCACCAGTCAAGAAAAGCATCTTACGGATGSGCATGACAGTAAGAGAAAT	2146
Qy	4091	tatgcagtgtgccataaccatgagtataaacactgcggcccaacttacttctgacaacga	4150
Db	2147	TATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA	2206
Qy	4151	tcggaggaccgaaggagctaaccgcttttttgcaacaacatggggggatcatgtaaactcgcc	4210
Db	2207	TCCGAGACCCGAAGGAGCTAACCCGCTTTTTCACAACATGGGGGATCATGTAACTCGCC	2266
Qy	4211	ttgatcgtttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacga	4270
Db	2267	TTGATCGTTGGGAACCCGGAGCTGAATGAAGCCATACCCAAACGAGCGGTGACACCCACGA	2326
Qy	4271	tgcctgtagcaaatggcaacaacgttgcgcaaacattataactggggaactacttactctag	4330
Db	2327	TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG	2386
Qy	4331	cttcccggaacaataatagactggatggaggcggaataaaagtgcaggaccacttctgc	4390
Db	2387	CTTCCCGGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACCTTCTGC	2446
Qy	4391	gctcggcccttccggctggctggtttattgtctgataaaatctggagccgggtgagcgtgggt	4450

Db	2447	GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCCGGTGAGCGTGGGT	2506
Qy	4451	ctcgcggtatcatcattgcagcactggggccagatggtgaagccctcccgatatcgtagttatct	4510
Db	2507	CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT	2566
Qy	4511	acacgacgggggagtcaggcaactatggatgaacgaataagacagatcgctgagatagggtg	4570
Db	2567	ACACGACGGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGATAGGTG	2626
Qy	4571	cctcactgattaaagcattggtaactgtcagaccaagtttactatataataacttttagattg	4630
Db	2627	CCTCACTGATTAAAGCATTGGTAACCTGTACAGACCAAGTTTACTCATATATACTTTAGATTG	2686
Qy	4631	atttaaaacttcatttttaaatttaaaggatctagtggaagatcccttttttgataaatctca	4690
Db	2687	ATTTAAAACTTCATTTTAAATTTAAAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA	2746
Qy	4691	tgaccaaaatcccttaacgtgagttttcgttcocactgagcgtcagaccccgtagaaaaaga	4750
Db	2747	TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA	2806
Qy	4751	tcaaaggatcttcttgagatcccttttttctgcgcgtaaatctgctgcttgcaaacaaaaa	4810
Db	2807	TCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGGTAATCTGCTGCTTGCAAAACAAAA	2866
Qy	4811	aaccacccgtaccagcgggtggtttgttcgccggtacaccaaatactgtccttctagtagccgtagt	4870
Db	2867	AACCACCGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGA	2926
Qy	4871	aggtaaactggcttcagcagagcgcagataaccacaaatactgtccttctagtagccgtagt	4930
Db	2927	AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGT	2986
Qy	4931	taggccacacttcaagaactctgtagcaccgcctacatacctcgtcctgtctaactcctgt	4990
Db	2987	TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT	3046
Qy	4991	taccagtggctgctgccagtggcgataaagtctgtcttaccgggttggaactcaagacgat	5050
Db	3047	TACCAGTGGCTGCTGCCAGTGCCGATAAGTCGTGTCTTACCAGGTTGGACTCAAGACGAT	3106
Qy	5051	agttaccgggataaaggcgcagcggctcgggtgaacgggggttcgtgcacacagcccagct	5110
Db	3107	AGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTCACACAGCCACGCT	3166
Qy	5111	tgagcgaacgacctacaccgaactgagataacctacagcgtgagctatgagaaagcgcca	5170
Db	3167	TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA	3226
Qy	5171	cgcttccgaaggaggaagggcggaaggtatccgggtacggcggaagcgtcggaacagggag	5230
Db	3227	CGCTTCCGAAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGTCGGAACAGGAG	3286
Qy	5231	agcgcacgagggagcttccagggggaaaacgcctggatatctttatagtcctgtcggttttc	5290
Db	3287	AGCGCACGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCTCTCGGGTTTC	3346
Qy	5291	gccacctctgacttgagcgtcgatttttgcgataacgcctggatatctttatagtcctgtcggttttc	5350
Db	3347	GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA	3406
Qy	5351	aaaaacgcagcaacgcggccctttttacgggttccctggccttttgccttttgctcaca	5410
Db	3407	AAAACGCCAGCAACGCGGCCCTTTTACGGTTCCTGGCCTTTTGTGCTGGCCTTTTGCTCACA	3466
Qy	5411	tgtcttctctgcttatcccttgattctgtggataaccggtattaccgcttttgagtgag	5470
Db	3467	TGTTCTTTCTCGCTTATCCCTTGATTTCTGTGGATAACCGTATTACCGCCTTTTGAGTGAG	3526
Qy	5471	ctgataccgctcgcgcgcagccgaacgcagcagcgcagcgcagtcagtcagtcagcgaagcgg	5530
Db	3527	CTGATACCGCTCGCCCGCAGCCCGAACGACCGAGCGCAGCGAGTCACTGAGCGGAGGACGG	3586

```
Qy 5531 aagagcgcccaatacgcaaacccctctctcccgcgcttgccgagcttaataatgcagct 5590
|||||
Db 3587 AAGAGCGCCAAATACGAAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCT 3646

Qy 5591 ggcacgacaggttcccgactggaaagcgggcagtgagcgcaacgcaataatgtgagtt 5650
|||||
Db 3647 GGCACGACAGGTTCCCGACTGGAAAGCGGCAGTGAGCGCAACGCAATTAATGTGAGTT 3706

Qy 5651 agctcactcattaggcacccccaggctttacacatttatgcttccgctcgatatgttggtg 5710
|||||
Db 3707 AGCTCACTCATTAGGCACCCCGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 3766

Qy 5711 gaattgtgagcgggataacaatttcacacaggaacacagctatgaccatgattacgcca 5767
|||||
Db 3767 GAATTGTGAGCGGATAACAATTTTCACACAGGAACAGCTATGACCATGATTACGCCA 3823

RESULT 3
UCU89964
LOCUS UCU89964 4473 bp DNA circular SYN 03-SEP-1997
DEFINITION Cloning vector pG5CAT, complete sequence.
ACCESSION U89964
VERSION U89964.1 GI:2293130
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Holtz,A. and Lou,Y.
TITLE pG5CAT complete sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4473)
AUTHORS Holtz,A. and Lou,Y.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-1997) CLONTECH Laboratories, Inc., 1020 East
Meadow Circle, Palo Alto, CA 94303-4230, USA
COMMENT This vector can be obtained from CLONTECH Laboratories, Inc., 1020
East Meadow Circle, Palo Alto, CA 94303-4230, USA. To place an
order call (415) 424-8222 or (800) 662-2566, extension 1.
International customers, please contact your local distributor. For
technical information, call (415) 424-8222 or (800) 662-2566,
extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH. If you suspect there is an
error in this sequence, please contact CLONTECH's Technical Support
Department at (415) 424-8222 or (800) 662-2566, extension 3 or
E-mail TECH@CLONTECH.COM.

FEATURES
source Location/Qualifiers
1. .4473
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
120. .779
/gene="CAT"
120. .779
/gene="CAT"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="AAB68658.1"
/db_xref="GI:2293131"
/translation="MEKKITGYTTVDISQWHRKEHFEAFQSVAAQCTYNQTVOLDITAF
LKTVKKNKHKFPYPAFIHILARLMAHPEFRMAKDGELVWDSVHPCYTVFHEQTETF
SLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGFLENMFVFSANPWVSFTSFDLNV
ANMNDFFAPVFTMGKYVTQGDVKVLMPLAIQVHHAVCDGFHVGRMLNELQYCDWEQGG
A"
2337. .3197
/gene="ampR"
2337. .3197
/gene="ampR"
/function="ampicillin resistance"
/codon_start=1
```

```
/transl_table=11
/protein_id="AAB68659.1"
/db_xref="GI:2293132"
/translation="MSIQHERVALIPFFAAFCLEPFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILESFRPEERFPMSTFKVLLCGAVLSRIDAGQEQLGRRIRHYSONDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLLHNMGDHVTRL
DRWPELNEAIPNDERDTPVAMATTLRKLLLTGELLTLASRQQLIDWMEADKVGAPL
LRSALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSGQATMDERNRQIA
EIGASLIKHW"
BASE COUNT 1183 a 1044 c 1046 g 1200 t
ORIGIN
```

```
Query Match 45.7%; Score 2637; DB 12; Length 4473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3131 cgaattcactggccgctgttttacaaagctgtgactgggaaaaacccctgtaccccaac 3190
|||||
Db 1735 CGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAACCCCTGGCGTTACCCCAAC 1794

Qy 3191 ttaatcgcttgacgacacatccccctttccgagctggcgtaataagcgaagagcccgca 3250
|||||
Db 1795 TTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 1854

Qy 3251 ccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatgcggtatt 3310
|||||
Db 1855 CCGATCGCCCTTCCCAACAGTTGGCGAGCCTGAATGGCGAATGGCGCCTGATGCGGTATT 1914

Qy 3311 ttctccttaacgcatctgtgcggtattttcacaccgcgcataatggtgcactctcagtaacaatct 3370
|||||
Db 1915 TTCTCCTTACGCATCTGTGCGGTATTTACACACCGCATATGGTGCATCTCAGTACAATCT 1974

Qy 3371 gctctgatgcgcgcatagttaagccagcccgacccccgcacaccccgctgacgcgcct 3430
|||||
Db 1975 GCTCTGATCGGCATAGTTAAGCCAGCCCGGACACCCGCCAACACCCGCTGACGCGCCCT 2034

Qy 3431 gacgggcttgctgctcccgcatcccgcttacagacaagctgtgacgctctccgggagct 3490
|||||
Db 2035 GACGGGCTTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 2094

Qy 3491 gcatgtgcagaggttttcaccgtcatcacggaacgcgcgagacgaaggcgctcgtga 3550
|||||
Db 2095 GCATGTGTCAAGAGGTTTTCACCGTCAATCACCGAAACGCGGAGACGAAAGGGCCTCGTGA 2154

Qy 3551 tacgcctattttatagggttaatgtcatgataataatgggtttcttagacgtcaggtggca 3610
|||||
Db 2155 TACGCCCTATTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTACACGCTCAGGTGGCA 2214

Qy 3611 cttttcggggaaatgtgcgcgggaacccctattttgtttatttttctaataacattcaata 3670
|||||
Db 2215 CTTTTCGGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATA 2274

Qy 3671 tgtatccgctcatgagacaataaacctgataaaatgcttcaataataattgaaaaaggaaga 3730
|||||
Db 2275 TGTATCCGCTCATGAGACAATAACCCCTGATAAAATGCTTCAATAATATGAAAAAGGAAGA 2334

Qy 3731 gtatgagtattcaacatttcgctgtgcgccttattcccttttttgcggcattttgccttc 3790
|||||
Db 2335 GTATGAGTATTCAAACATTTCGCTGTGCGCCTTATTCCTTTTTCGGGCATTTTGCCCTTC 2394

Qy 3791 ctgtttttgctcaccagaaacgcgtggtgaaagtataaagatgctgaagatcagttgggtg 3850
|||||
Db 2395 CTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAGATCAGTTGGGTG 2454

Qy 3851 cacgagtggttacatcgaaactggatctcaacacgcggtaagatcccttgagagttttcgcc 3910
|||||
Db 2455 CACGAGTGGGTTACATCGAACTGGATCTCAACACGGGTAAAGATCCCTTGAGAGTTCCTCGCC 2514

Qy 3911 ccgaagaacggttttccaatgatgagcacttttaaagtctctgtatgtgcgcggtattat 3970
|||||
Db 2515 CCGAAGAACGTTTTCCAATGATGAGCACTTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 2574
```


QY 3971 cccgtattgacgccgggcaagagcaaacctcggtcgccgcatacacactattctcagaatgact 4030
|||||
Db 2575 CCCGTATTGACCCGGGCAAGACAACTCGGTGCGCGCATACACTATTCTCAGAATGACT 2634
QY 4031 tgggtgagtgactcaccagtcacagaaaaagcatctttacggatggcatgacagtaagagaat 4090
|||||
Db 2635 TGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT 2694
QY 4091 tatgcagtgctgccataaacctagtgatgataaacactgcggccaacttacttctgacaacga 4150
|||||
Db 2695 TATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA 2754
QY 4151 tcggaggaaccgaagagctaaccgcttttttgacaaacatggggggatcatgttaactcgcc 4210
|||||
Db 2755 TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCC 2814
QY 4211 ttgatcgttgggaaccggagctgaatgaagccataccaaacagcagcgtgacaccacga 4270
|||||
Db 2815 TTGATCGTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA 2874
QY 4271 tgcctgtagcaatggcaacaacgttgcgcaaacattataactggcgaaactacttactctag 4330
|||||
Db 2875 TGCTGTAGCAATGGCAACAACGTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 2934
QY 4331 ctccccggcaacaattaatagactggatggaggcgggataaaagtgcaggaccacttctgc 4390
|||||
Db 2935 CTTCGCCGGCAACAATTAAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACCTCTGC 2994
QY 4391 gctcggcccttccggctggctggtttattgtctgataaaatctggagccggtgagcgtgggt 4450
|||||
Db 2995 GCTCGGCCCTTCCGGCTGGCTGTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT 3054
QY 4451 ctcggggtatcattgcagcaactggggccagatggttaagccctcccgtatcgtagttatct 4510
|||||
Db 3055 CTCGGGATATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 3114
QY 4511 acacgacggggagtcaggcaactatggtgaacgaaaaatagacagatcgctgagatagtg 4570
|||||
Db 3115 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGTG 3174
QY 4571 cctcaactgattaagcattggttaactgtcagaccaaagtttactcatatatacttttagattg 4630
|||||
Db 3175 CCTCACTGATTAAAGCATTTGGTAACGTGTCAGACCAAGTTTACTCATATATATACTTTAGATTG 3234
QY 4631 atttaaaaactcatttttaattaaaaggatctaggtgaagatcccttttggataaatctca 4690
|||||
Db 3235 ATTTAAAACTTCATTTTAAATTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 3294
QY 4691 tgaccaaaaatcccttaacgtgagtttctggtccactgagcgtcagacccccgtagaaaaa 4750
|||||
Db 3295 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA 3354
QY 4751 tcaaaaggatcttcttgagatcccttttctgcgcgttaactgctgcttgcaaaacaaaa 4810
|||||
Db 3355 TCAAAGGATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAA 3414
QY 4811 aaccaccgctaccagcgggtgtgtgttgcgggatcaagagctaccaactcttttccga 4870
|||||
Db 3415 AACCACCGCTACCAGCGGTGGTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGA 3474
QY 4871 agttaactggcttcagcagagcgagataccaaaactgtccttctagtgtagccgtagt 4930
|||||
Db 3475 AGCTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGCTCTTAGTGAGCCCGTAGT 3534
QY 4931 taggccaccacttcaagaactctgtagcaccgcctacataacctcgtctgctaactcgt 4990
|||||
Db 3535 TAGGCCACCACCTTCAAGAACTCTGTAGCACCCGCCCTACATACTCGCTCTGCTAATCCTGT 3594
QY 4991 taccagtggtcgtcgtccagtgccgataaagtcgtgtcttaccgggttggaactcaagacgat 5050
|||||
Db 3595 TACCAGTGGCTGCTGCCAGTGGCGATAAAGTCGTGTCTTACCGGGTTGGACTCAAGACGAT 3654
QY 5051 agttaccgggataaggcgacggtcggtggtgaacgggggttcgtgcacacagccccagct 5110

Db 3655 AGTTACCGGATAAGCGCAGCGTGGGCTGAACGGGGGTTCGTGCACACAGCCCAGCT 3714
QY 5111 tggagcgaacgacctacacccgaactgagatacctacagcgtgagctatgagaaagcgcca 5170
|||||
Db 3715 TGGAGCGAAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 3774
QY 5171 cgctcccgaaaggagaaaggcgacaggtatccggtaagcgagggcggaacaggag 5230
|||||
Db 3775 CGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGGGTCGGAACAGGAG 3834
QY 5231 agcgacgaggagagcttccagggggaaaaacgcctggtatctttatagtcctgtcgggttcc 5290
|||||
Db 3835 AGCGACGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCCTGTCTGGGTTTC 3894
QY 5291 gccacctctgacttgagcgtcgatttttctgctgctcgtcagggggcgagcctatgga 5350
|||||
Db 3895 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA 3954
QY 5351 aaaacgcccagcaacgcggccctttttacgggttcccttgcccttttgctggccttttgcacaa 5410
|||||
Db 3955 AAAACGCCAGCAACGCGGCCCTTTTACGGTTCCTGGCCCTTTTGTGGCTTGTGTCACA 4014
QY 5411 tgttcttctcgtgttatccccctgattctgtggataaccgtattaccgccttttgagtga 5470
|||||
Db 4015 TGTCTTCTCCTGCGTTATCCCTGATTTCTGGATAACCGTATTACCGCCTTTTGAGTGAG 4074
QY 5471 ctgataccgctcgccgcagccgaacgaccgagcgcagcagtcagtcagtcaggaagcgg 5530
|||||
Db 4075 CTGATACCGCTCGCCGCAGCCGACGACCGGAGCGCAGTCAGTCAGGAGGAGCGG 4134
QY 5531 aagcgcgccccaaatacgcaaacccctctccccgcgcgttgccgagttcattaaatgcagct 5590
|||||
Db 4135 AAGAGCGCCCAATACGCAAACCGCTCTCCCCGCGTGGCCGATTTCATTAAATGCAGCT 4194
QY 5591 ggcagcagaggtttccccgactggaagcgggagtcagtcagtcagcaacgcaatgaatgtgagtt 5650
|||||
Db 4195 GGCACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 4254
QY 5651 agctcaactcattagggcaccgccaggttttacacctttatgcttccggctcgtatgtgtg 5710
|||||
Db 4255 AGCTCACTCATTAGGCACCCCGAGGCTTACACTTTATGCTTCGGCTCGTATGTTGTGTG 4314
QY 5711 gaattgtgagcgggataaacaatttcacacaggaacagctatgacctgattacgcca 5767
|||||
Db 4315 GAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCCA 4371

RESULT 4

AX191663

LOCUS AX191663

DEFINITION Sequence 29 from Patent WO0149832.

ACCESSION AX191663

VERSION AX191663.1 GI:15209844

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .4960

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="vector pCMV-I-Cre-pA"

BASE COUNT 1225 a 1213 c 1296 g 1226 t

ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 4960;

Best Local Similarity 100.0%; Pred. No. 0;				
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	3131	cgaattcaactgcccgtcgtttttacaacgctcgtgactgggaaaaacccctggcttaccccaac	3190	
Db	2289	CGAATTCACTGGCCGTCGTTTACACACGTCGTGACTGGGAAAACCCCTGGCGTTACCCCAAC	2348	
Qy	3191	ttaatcgccctgcagcacatccccctttccgcagctggcgtaataagcgaagagggcccgca	3250	
Db	2349	TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA	2408	
Qy	3251	ccgatcgcccttcccaacagttgcgagccctgaatggcgaaatggcgccctgatcggtattt	3310	
Db	2409	CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTGATCGGGTATT	2468	
Qy	3311	ttctccttaocgatctgtgcggtatttcaacaccgcataatggtgaactctcagtacaatct	3370	
Db	2469	TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACCTCTCAGTACAAATCT	2528	
Qy	3371	gctctgatcgcgcatagttaagccagcccgacacccgcgaaccccgctgacgcgcct	3430	
Db	2529	GCTCTGATGCGGCATAGTTAAGCCAGCCCGCACACCCGCCCAACCCGCTGACCGCCCT	2588	
Qy	3431	gacgggctgtctgctcccggcacccgttacagacaagctgtgacgctctccgggagct	3490	
Db	2589	GACGGGCTTGTCTCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT	2648	
Qy	3491	gcatgtgtcagaggttttcacccgtcatcacccgaaacgcgcgagagcaaaaggccctcgtga	3550	
Db	2649	GCATGTGTGAGAGGTTTTCACCGTCTACCCGAACCGCGGAGACGAAGGGCCCTCGTGA	2708	
Qy	3551	tacgcctatttttataggttaatgtcatgataataatggtttcttaaatacattcaaatga	3610	
Db	2709	TACGCCTATTTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTAGACGTCAGGTGGCA	2768	
Qy	3611	cttttcggggaaaatgtgcgcggaacccctatttgtttatttttctaaatacatctcaaatga	3670	
Db	2769	CTTTTCGGGAAATGTGCGGGAACCCCTATTGTGTTATTTCATAATACATTCAATA	2828	
Qy	3671	tgtatccgctcatgagacaataacccctgataaaatgcttcaataataattgaaaaaggaaga	3730	
Db	2829	TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCATAATATGMAAAGGAAGA	2888	
Qy	3731	gtagagtattcaacatttcctgctgcgcccttattcccttttttgcggcattttgccttc	3790	
Db	2889	GTATGAGTATCAACATTTCCGTTGTCGCCCTTATTCCCCTTTTTCGGGCATTTTGCCCTTC	2948	
Qy	3791	ctgtttttgctcaccagaaaacgctggtgaaaagtaaaagatgctgaagatcagttgggtg	3850	
Db	2949	CTGTTTTTGTCACCCCAGAACCGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG	3008	
Qy	3851	cacgagtggttacatcgaaactggatcacaacagcgggtaagatcccttgagagttttcgcc	3910	
Db	3009	CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCC	3068	
Qy	3911	ccgaagaacgttttccaatgatgagcaacttttaaagttctgctatgtggcgcggtattat	3970	
Db	3069	CCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT	3128	
Qy	3971	cccgtattgacgcggggcaagagcaactcggctgcgcgcatacacatttctcagaatgact	4030	
Db	3129	CCCCGATTGACGCGGGCAAGAGCACTCGGTCGCCGCATACACTATTCTCAGAATGACT	3188	
Qy	4031	tggttgagtactcacccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat	4090	
Db	3189	TGGTTGAGTACTCACCATGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA	3248	
Qy	4091	tatgcagtctgcccataaccatgagtataaacactgcggccaacttacttctgacaacga	4150	
Db	3249	TATGCAGTCTGCCATAACCATGAGTGATAACACTGCGGCCCACTTACTTCTGACAACGA	3308	
Qy	4151	tcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgcc	4210	

Db	3309	TCGGAGGACCGAAGGAGCTAACCCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCC	3368	
Qy	4211	ttgatcgttgggaaaccggagctgaatgaagccataccaaaacgagcgtgacacccacga	4270	
Db	3369	TTGATCGTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA	3428	
Qy	4271	tgccgttagcaaatggcaacaacgttgcgcaaaactattaactggcgaactacttactctag	4330	
Db	3429	TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG	3488	
Qy	4331	cttcccggaacaataatagactggatggaggcgggataaaagtgcaggaccacttctgc	4390	
Db	3489	CTTCCCGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACCTTCTGC	3548	
Qy	4391	gctcggcccttccggctggctggttttattgtctgataaaatcctggagcggtagcgtgggt	4450	
Db	3549	GCTCGGCCCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCGGTGAGCGTGGGT	3608	
Qy	4451	ctcgcggtatcatctgcagcactggggccagatggtaagccctcccgatatcgttagttatct	4510	
Db	3609	CTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCCTATCGTAGTTATCT	3668	
Qy	4511	acacgacggggagtcaggcaactatgatgaacgaaaaatagacagatcgctgagataggtg	4570	
Db	3669	ACACGACGGGAGTCAAGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTG	3728	
Qy	4571	cctcactgattaagcattggtaaactgtcagaccaaagtttactactatatactttagattg	4630	
Db	3729	CCTCACTGATTAAAGCATTTGTAACCTGTACAGACCAAGTTTACTCATATATACTTTAGATTG	3788	
Qy	4631	atttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttggataatctca	4690	
Db	3789	ATTATAAACTTCATTTTAAATTAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCA	3848	
Qy	4691	tgaccaaaaatcccttaacgtgagtttttcgttccactgagcgtcagacccccgtagaaaaa	4750	
Db	3849	TGACCAAAATCCCTTAACGTGAGTTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA	3908	
Qy	4751	tcaaggatcttcttgagatccctttttctgcgcgtaaatctgctgcttgcaaaacaaaa	4810	
Db	3909	TCAAAGGATCTTCTTGAGATCCCTTTTTTCTGCGCGTAATCTGCTGCTGCAAAACAAAA	3968	
Qy	4811	aaccacgcctaccagcgggtggttbtgttcggcgatcaagagctaccaactcttttccga	4870	
Db	3969	AACCACCGCTACCAGCGGTGGTTTGTTCGCCGGATCAAGAGCTACCAACTCTTTTCCGA	4028	
Qy	4871	agttaactggcttcagcagcgcagataccaaaatactgccttcttagtgtagccgtagt	4930	
Db	4029	AGGTAAC TGSGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGTGTAGCCGTAGT	4088	
Qy	4931	taggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgtctaactcctgt	4990	
Db	4089	TAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCCTGT	4148	
Qy	4991	taccagtgctgctgccagtgccgataaagtgcgtgtcttaccgggttggactcaagacgat	5050	
Db	4149	TACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTTACCGGGTTGGACTCAAGACGAT	4208	
Qy	5051	agttaccggataaaggcgcagcgggtcgggctgaacgggggttcctgcacacagccccagct	5110	
Db	4209	AGTTACCGGATAAAGGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTCACACACAGCCAGCT	4268	
Qy	5111	tggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaaagcgcca	5170	
Db	4269	TGGAGCGAACCACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCGCCA	4328	
Qy	5171	cgcttccgaaggagaaaagcggcacaggtatccggtaagcggcaggggtcggaacaggag	5230	
Db	4329	CGCTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCGGTCGGAACAGGAG	4388	
Qy	5231	agcgcacgagggagcttccaggggggaacgcctggatatctttatagtcctgtcggggttc	5290	
Db	4389	AGCGCACGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCTCTGTCGGGTTC	4448	

QY	5291	gccacctctgacttgagctcgatcttctgtgatgctcgtcaggggggagcctatgga	5350
Dd	4449	gccacctctgacttgagcgtcgatcttctgtgatgctcgtcaggggggagcctatgga	4508
QY	5351	aaaacgccagcaacgcgccttttttacggttcctggccttttgctgctcaca	5410
Dd	4509	aaaacgccagcaacgcgccttttttacggttcctggccttttgctgctcaca	4568
QY	5411	tggtcttctcgttatccctgattctgtggataaacctattaccgccttgagtga	5470
Dd	4569	tggtcttctcgttatccctgattctgtggataaacctattaccgccttgagtga	4628
QY	5471	ctgataccgctcgcgcagcgaacagcagcagcagcagcagcagcagcagcagc	5530
Dd	4629	ctgataccgctcgcgcagcgaacagcagcagcagcagcagcagcagcagcagc	4688
QY	5531	aagagcgcccaatacagcaaacgcgcctctccccgcgcgttgccgattcattaatgcagct	5590
Dd	4689	aagagcgcccaatacagcaaacgcgcctctccccgcgcgttgccgattcattaatgcagct	4748
QY	5591	ggcagcagcaggtttcccgactggaaagcgggagcagcagcagcagcagcagcagc	5650
Dd	4749	ggcagcagcaggtttcccgactggaaagcgggagcagcagcagcagcagcagcagc	4808
QY	5651	agctcactcattaggcaccccgctttacacttttatgcttcggcgtcgtatgtgtg	5710
Dd	4809	agctcactcattaggcaccccgctttacacttttatgcttcggcgtcgtatgtgtg	4868
QY	5711	gaattgtgagcggataaacaatttcacacaggaacagctatgacctgattacgcca	5767
Dd	4869	gaattgtgagcggataaacaatttcacacaggaacagctatgacctgattacgcca	4925
RESULT	5		
AX114861	AX114861	5711 bp	DNA
LOCUS	Sequence 21 from Patent WO0129208.	linear	PAT 11-MAY-2001
DEFINITION	AX114861		
ACCESSION	AX114861		
VERSION	AX114861.1	GI:14031803	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	synthetic construct		
REFERENCE	artificial sequence.		
AUTHORS	1 (bases 1 to 5711)		
TITLE	Kuehn,R., von Melchener,H. and Altschmied,J.		
JOURNAL	Conditional gene trapping construct for the disruption of genes		
FEATURES	Patent: WO 0129208-A 21 26-APR-2001;		
source	ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)		
BASE COUNT	Location/Qualifiers		
ORIGIN	1. .5711		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="vector PRK65"		
	1329 a 1511 c 1617 g 1254 t		
Query Match	45.7%;	Score 2637;	DB 6; Length 5711;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2637;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	3131	cgaattcactggcgcgtctgttttacaacgctcgtgactgggaacccctggcgttacccaac	3190
Dd	3040	cgaattcactggcgcgtctgttttacaacgctcgtgactgggaacccctggcgttacccaac	3099
QY	3191	ttaatgccttgacagcacatccccctttcgccagctggcgtaataagcgaagggcccgca	3250
Dd	3100	ttaatgccttgacagcacatccccctttcgccagctggcgtaataagcgaagggcccgca	3159
QY	3251	ccgatgccttcccaacagttggcagcctgaatggcggaatggcgctgatgcggtatt	3310
Dd	3160	ccgatgccttcccaacagttggcagcctgaatggcggaatggcgctgatgcggtatt	3219

QY	3311	ttctccttacgcacatctgtgcggtattttcacacaccgcacatatggtgcactctcagtacaaatct	3370
Dd	3220	ttctccttacgcacatctgtgcggtattttcacacaccgcacatatggtgcactctcagtacaaatct	3279
QY	3371	gctctgatcccgcatagttaagcagcccccgacacccccgaacaccccgctgacgcgcct	3430
Dd	3280	gctctgatcccgcatagttaagcagcccccgacacccccgaacaccccgctgacgcgcct	3339
QY	3431	gacgggcttctgctctccccgcacccgttacagacaagctgtagccgctctccgggagct	3490
Dd	3340	gacgggcttctgctctccccgcacccgttacagacaagctgtagccgctctccgggagct	3399
QY	3491	gcatgtgtcagaggtttttcacccgtcatcacgaaacgcgcgagacgaaaggcctcgtga	3550
Dd	3400	gcatgtgtcagaggtttttcacccgtcatcacgaaacgcgcgagacgaaaggcctcgtga	3459
QY	3551	tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcagggtgca	3610
Dd	3460	tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcagggtgca	3519
QY	3611	ctttcgggggaaatgtgcgggaaacccctatttgtttattttctaaatacattcaata	3670
Dd	3520	ctttcgggggaaatgtgcgggaaacccctatttgtttattttctaaatacattcaata	3579
QY	3671	tgatcgcgtcatgagacaataaacctgataaaatgcttcaataatatgaaaaaggaaga	3730
Dd	3580	tgatcgcgtcatgagacaataaacctgataaaatgcttcaataatatgaaaaaggaaga	3639
QY	3731	gtatgagtattcaacatttccgtgtgcgccttatttcccttttttggcgcattttgccttc	3790
Dd	3640	gtatgagtattcaacatttccgtgtgcgccttatttcccttttttggcgcattttgccttc	3699
QY	3791	ctgttttctcaccacccagaaacgctgggtgaaagtaaaagatgctgaagatcagttgggtg	3850
Dd	3700	ctgttttctcaccacccagaaacgctgggtgaaagtaaaagatgctgaagatcagttgggtg	3759
QY	3851	cacgagtgggttacatcgaaactgagatctcaacagcggtaagatccttgagagtttctgcc	3910
Dd	3760	cacgagtgggttacatcgaaactgagatctcaacagcggtaagatccttgagagtttctgcc	3819
QY	3911	ccgaagaacgcttttccaatgatgagcacttttaagttctgctatgtggcggtattat	3970
Dd	3820	ccgaagaacgcttttccaatgatgagcacttttaagttctgctatgtggcggtattat	3879
QY	3971	ccggtattgacgcgggcaagagcaactcggcgccgcacatacactattctcagaatgact	4030
Dd	3880	ccggtattgacgcgggcaagagcaactcggcgccgcacatacactattctcagaatgact	3939
QY	4031	tggttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat	4090
Dd	3940	tggttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat	3999
QY	4091	tatgcagtgtgcccataaaccatgagtgataaacactgcggccaaacttacttctgacaaacga	4150
Dd	4000	tatgcagtgtgcccataaaccatgagtgataaacactgcggccaaacttacttctgacaaacga	4059
QY	4151	tcggaggaccgaaggagctaacccgttttttgcaacaacatgggggatcatgtaactcgcc	4210
Dd	4060	tcggaggaccgaaggagctaacccgttttttgcaacaacatgggggatcatgtaactcgcc	4119
QY	4211	ttgatcgttgggaacccggagctgaatgaagccataccaaaacgacgagcgtgacaccacga	4270
Dd	4120	ttgatcgttgggaacccggagctgaatgaagccataccaaaacgacgagcgtgacaccacga	4179
QY	4271	tgccctgtagcaatggcaacaacgttggcgaacactattaactggcgaactacttactctag	4330
Dd	4180	tgccctgtagcaatggcaacaacgttggcgaacactattaactggcgaactacttactctag	4239
QY	4331	cttcccgccgaacaataatagactggatggagcggtataaagtgcaggaccacttctgc	4390
Dd	4240	cttcccgccgaacaataatagactggatggagcggtataaagtgcaggaccacttctgc	4299

QY 4391 gctcgccctccggtgctggtttattgctgataaaatctggagccggtgagcgtgggt 4450
|||||
Db 4300 GCTCGGCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGACGTTGGGT 4359
QY 4451 ctcgcggtatcattgcagcactggggccagatggtgaagccctcccgatatcgtagtatatc 4510
|||||
Db 4360 CTCGGGTATCATTCAGCACCTGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 4419
QY 4511 acacgacggggagtcaggcaactatgtagaacgaaatagacagatcgctgagatagtg 4570
|||||
Db 4420 ACACGACGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTG 4479
QY 4571 cctcactgattaagcattggttaactgacagaccaagtcttactcatatatactttagattg 4630
|||||
Db 4480 CCTCACTGATTAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTG 4539
QY 4631 atttaaaacttcattttttaatttaaaggatcctaggtgaagatccttttttgataatctca 4690
|||||
Db 4540 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 4599
QY 4691 tgaccaaaatcccttaacgtgagtttctggtccactgagcgtcagacccccgtagaaaaaga 4750
|||||
Db 4600 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAAGACCCCGTAGAAAAAGA 4659
QY 4751 tcaaaggatctcttgagatccctttttctgcgtaaatctgctgcttagtagccgtagt 4810
Db 4660 TCAAAGGATCTCTTTGAGATCCTTTTCTGCGGTAATCTGCTGCTTGCATAACAAAAA 4719
QY 4811 aaccacgcctaccagcggtggtgttctgcccggatcaagagctaccaactcttttccga 4870
|||||
Db 4720 AACCACCGCTACCAGCGGTGTTGTTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGA 4779
QY 4871 aggtaactggttcagcagcgagcgagataccaaaatactgtcctcttagtagtagccgtagt 4930
|||||
Db 4780 AGGTAACCTGCTTCAGCAGAGCGGCAGATACCAATACTGTCCCTTCTAGTGTAGCCGTAGT 4839
QY 4931 taggccaccacttcaagaactctgtagcaccgcctacatacacctcgtctgtaactcctgt 4990
|||||
Db 4840 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGTCTGTCTAATCCTGT 4899
QY 4991 taccagtgctgctgccagtgccgataagtcgctgtcttaccgggttggaactcaagacgat 5050
|||||
Db 4900 TACCAGTGGCTGCTGCCAGTGCCGATAAGTCGTGCTTTACCGGGTTGGACTCAAGACGAT 4959
QY 5051 agttaccgataaaggcgagcagcgtcggtcggtgaaacgggggttcgtgacacagcccagct 5110
|||||
Db 4960 AGTTACCGGATAAAGCGCAGCGGTCCGGCTGAACGGGGGTTCCGTGCACACAGCCCAGCT 5019
QY 5111 tggagcgaacgacctacacccaactgagatacctacagcgtgagctatgagaaagcgcca 5170
|||||
Db 5020 TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 5079
QY 5171 cgcttcccgaaggagagaaaggcggaaggtatccggtgaagcggcagggtcggaacaggag 5230
|||||
Db 5080 CGCTTCCCGAAGGGAGAAAGCGGCACAGGTATCCGGTAAGCGGCAGSGTCGGAACAGGAG 5139
QY 5231 agcgcacgagggagcttccaggggggaaacgcctggtatctttatagtcctgctcggtttc 5290
|||||
Db 5140 AGCGCACGAGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTTATAGTCTGTCGGGTTTC 5199
QY 5291 gccacctctgacttgagcgtcgatcttttgatgctcgtcagggggcggaagcctatgga 5350
|||||
Db 5200 GCCACCTCTGACTTGAGCGTCGATTTTGTGTATGCTCGTCAGGGGGGCGGAGCCCTATGGA 5259
QY 5351 aaaacgccagaaacgcggccctttttaagggttccttgcccttttgctggcccttttgctcaca 5410
|||||
Db 5260 AAAACGCCAGCAACGCGGCCCTTTTACGGTTCCTTGGCCCTTTTGTGCGCCTTTTGTCTACA 5319
QY 5411 tgttcttctcgtgttatccctgattctgtggataaccgtattacgccttttgagtgag 5470
|||||
Db 5320 TGTTCCTTCTGCGTTATCCCTTGATTCTGTGGATAACCCGTATTACCGCCTTTTGAGTGAG 5379
QY 5471 ctgataaccgctcgccgcagcccgaaacgacgagcgcagcagagtcagtgagcgaagcgg 5530

|||||
Db 5380 CTGATACCGTCGCCGCGAGCCGAACGACGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGG 5439
QY 5531 aagagcgcccaatacagcaaacccgcctctcccgcgctgtggccgattcattaatgcagct 5590
|||||
Db 5440 AAGAGCGCCCAATAGCAAAACCGCCTCTCCCGCGCTTGCCGATTCATTAAATGCAGCT 5499
QY 5591 ggcacgacaggtttcccgactggaaagcgggcagtgagcgcacaacattaatgtgagtt 5650
|||||
Db 5500 GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 5559
QY 5651 agctcactcattaggcacccccaggcctttacactttatgcttccggctcgtatgtgtg 5710
|||||
Db 5560 AGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 5619
QY 5711 gaattgtgagcggataaacaatttcacacaggaacacagctatgaccatgattacgcca 5767
|||||
Db 5620 GAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCA 5676

RESULT 6
AX191664
LOCUS AX191664 7332 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 30 from Patent WO0149832.
ACCESSION AX191664
VERSION AX191664.1 GI:15209845
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 7332)
AUTHORS Schwenk,F.
TITLE Transduction of recombinases for inducible gene targeting
JOURNAL Patent: WO 0149832-A 30 12-JUL-2001;
ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES
source
1. 7332
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pCMV-I-beta-pA"

BASE COUNT 1718 a 1895 c 1964 g 1755 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 7332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactggccgctcgttttacaacgtcgtagctgggaaacccctggttaccacaac 3190
|||||
Db 4661 CGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAACCCCTGGCGTTACCCCAAC 4720
QY 3191 ttaatcgcttgagcacacatccccctttcccgagctggcgtaatagagaagagcccgca 3250
|||||
Db 4721 TTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 4780
QY 3251 ccgatcgcccttcccaacagttgcgagcctgaaatggcgaatggcgctgatgcggtatt 3310
|||||
Db 4781 CCGATCGCCCTTCCCAACAGTTGCCGACGCTGAATGGCGAATGGCGCTGATCGCGTATT 4840
QY 3311 ttctccttagcgcattcgtgcgggtattttcacaccgcataatggtgcactctcagtaaatct 3370
|||||
Db 4841 TTCTCCTTACGCAATCTGTGCGGTATTTTACACCCGCATATGGTGCATCTCTCAGTACAATCT 4900
QY 3371 gctctgatgcgcgcatagtttaagccagccccgcagaccccccaacacccgctgacgcgccct 3430
|||||
Db 4901 GCTCTGATGCCGCATAGTTAAGCCAGCCCGCACACCCGCCAACACCCGCTGACGCGCCCT 4960
QY 3431 gacgggcttctgctcccgcatcccgcttacagacaagctgtgacgcgtctccgggagct 3490
|||||
Db 4961 GACGGGCTTGTCTGCTCCCGGCATCCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 5020
QY 3491 gcatgtgtcagagggttttcaccgtcatcacgaaacgcgcgagacgaaagggcctcgtga 3550

Db 5021 GCATGTGTCAGAGGTTTTACCGTTCATCACCAGAAACGCCGAGACGAAAGGCCCTCGTGA 5080
QY 3551 tacgcctatttttatagggttaatgtcatgataataatagggtttcttagacgtcaggtggca 3610
Db 5081 TACGCCATTTTTATAGGTTAATGTCAATGATAATAATAGGTTCTTAGACGTCAGGTGGCA 5140
QY 3611 cttttcgggggaaatgtgcgcggaaacccctatttgtttatttttctaaataacattcaaaata 3670
Db 5141 CTTTTCGGGGAATGTGCGGGAACCCCTATTGTGTTATTCTTAAATACATTCAAAATA 5200
QY 3671 tgtatccgtcatgagacaataaacctgataaaatgcttcaataatattgaaaaaggaaaga 3730
Db 5201 TGTATCCGCTCATGAGACAATAACCCCTGATAAAATGCTTCAATAATAATTGAAAAAGGAAGA 5260
QY 3731 gtatgagtattcaacatttcctgctgcgccttattcccttttttgcggcatttgccttc 3790
Db 5261 GTATGAGTATTC AACATTTCCTGTGCGCCCTTATTCCTTTTTCGCGCATTTTGCCCTTC 5320
QY 3791 ctgtttttgctcaccagaaacgcgtggtgaaagtaaaagatgctgaagatcagttgggtg 3850
Db 5321 CTGTTTTTGTCTACCCAGAAACGCTGGTGAAGATAAAGATGCTGAAGATCAGTTGGGTG 5380
QY 3851 cacgagtggttacatcgaaactggatctcaacagcggtaagatccttgagagttttcgcc 3910
Db 5381 CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCCTTGAGAGTTTTCGCC 5440
QY 3911 ccgaagaacgttttccaatgatgagcactttttaaagttctgctatgtggcgcggtattat 3970
Db 5441 CCGAAGAACGTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 5500
QY 3971 cccgtattgacgcggggcaagagcaactgggtgcgcgcatacactattctcagaatgact 4030
Db 5501 CCCGTATTGACGCGCGGCAAGAGCAACTCGTGC CGCATACACTATTCTCAGAATGACT 5560
QY 4031 tgggttgagtactcaccagtcacagaaaaagcatcttacggatggcatggcatgacagtaagaaat 4090
Db 5561 TGGTTGAGTACTCACCAGTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAT 5620
QY 4091 tatgcagtgctgcataaaccatgagtataaacactgcggccaaacttacttctgcacaacga 4150
Db 5621 TATGCAGTCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA 5680
QY 4151 tcggaggacccgaaggagctaaccgcttttttgcaaacatgggggatcoatgtaaactcgcc 4210
Db 5681 TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCC 5740
QY 4211 ttgatcgttggaacccggagctgaatgaagccataccaaacgacgagcgtgacaccacga 4270
Db 5741 TTGATCGTTGGAAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGA 5800
QY 4271 tgcctgtagaatggcaacaacgttgcgaaaactattaaactggcgaaactactactactag 4330
Db 5801 TGCCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 5860
QY 4331 cttcccgggcaacaattaatagactggatggagggcggaataaaagttgcaggaccacttctgc 4390
Db 5861 CTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGC 5920
QY 4391 gctcggcccttcggcgtggctggtttattgctgataaaatctggagccggtgagcgtgggt 4450
Db 5921 GCTCGGCCCTTC CGGCTGGCTGGTTTTATTGCTGATAAAATCTGGAGCCCGTGAGCGTGGGT 5980
QY 4451 ctgcggtatcattgcagcactggggccagatggtaagccctcccgatcgttagttatct 4510
Db 5981 CTCGCGGTATCATTCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6040
QY 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtcgagatagggtg 4570
Db 6041 ACACGACGGGGAGTCAGGCAACTATGGATGACGAATAGACAGATCGTGAATAGGTG 6100
QY 4571 cctcactgattaagcattggttaactgtcagaccaaagtttactcataataacttttagattg 4630
|||||

Db 6101 CCTCACTGATTAAAGCATTTGGTAACGTGTCAAGCAAGTTTACTCATATATACTTTAGATTG 6160
QY 4631 atttaaaacttcatttttaatttaaaaggatctaggtgaagatccctttttgataaatctca 4690
Db 6161 ATTTAAAAACTTCATTTTAAATTTAAAAAGGATCTAGGTGAAGATCCCTTTTGATAATCTCA 6220
QY 4691 tgaccaaaaaatcccttaaacgtgagtttttcgctccactgagcgtcagaccccgtagaaaaaga 4750
Db 6221 TGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA 6280
QY 4751 tcaaaaggatcttcttgagatcccttttttttgcgcgtaaatctgctgcttgcaaaacaaaa 4810
Db 6281 TCAAAGGATCTTCTTGAGATCCCTTTTTTCTGCGCGTAATCTGCTGCTTGC AAAA 6340
QY 4811 aaccaccgctaccagcgggtggtttgtttgcgggataccaaagctaccaactcttttccga 4870
Db 6341 AACCACCGTACAGCGGTGTTTGTGTTGCGGATCAAGAGCTACCAACTCTTTTCCGA 6400
QY 4871 aggtaaactggcttcagcagagcgagataccaaaaactactgctccttctagtagccgtagt 4930
Db 6401 AGGTAAC TGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTTAGTGTAGCCGTAGT 6460
QY 4931 taggccaccacttcaagaactctgtagcacgcgcctacatacctcgtcctgctaactcctgt 4990
Db 6461 TAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGTAATCCTGT 6520
QY 4991 taccagtggtgctgccagtgccagtgccgataaagtsgtgccttacccgggttggaactcaagacgat 5050
Db 6521 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTTACCGGGTTGGACTCAAGACGAT 6580
QY 5051 agttaccgggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagct 5110
Db 6581 AGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGTTTCGTGCACACAGCCAGCT 6640
QY 5111 tggagcgaacgacctacacccgaactgagatacctacagcgtgagctatgagaaagcgcca 5170
Db 6641 TGGAGCGAACGACTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 6700
QY 5171 cgcttcccgaaaggagaaaggcggacagggtatccggtaagcggcagggtcggaacaggag 5230
Db 6701 CGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAG 6760
QY 5231 agcgcacgaggggagcttccagggggaaacgcctgggtatctttatagtcctgctcggtttc 5290
Db 6761 AGCGCACGAGGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCCCTGTCGGGTTTC 6820
QY 5291 gccacctctgacttgagcgtcgatatttttgatgctcgtcagggggcgagccctatgga 5350
Db 6821 GCCACCTCTGACTTGACGCTCGATTTTTTGTGATGCTCGTCAAGGGGGCGGAGCCTATGGA 6880
QY 5351 aaaacgccagcaacgcggcctttttacgggttccctggccttttgctggccttttgctcaca 5410
Db 6881 AAAACGCCAGCAACGCGGCCCTTTTACGGTTCTCGGCCCTTTTGTGGCCTTTTGCTCACA 6940
QY 5411 tgttcttctcgtggttatccctgattctgttgataaaccgtattaccgcctttgagtgag 5470
Db 6941 TGTTCCTTTCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG 7000
QY 5471 ctgataccgctcgccgcagccgaacgacgagcgcagcagcagtcagtcagtcagcgaagcgg 5530
Db 7001 CTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGGAGTCAGTGAGCGAGGAAGCGG 7060
QY 5531 aagagcgcceaataacgcaaacccctctctcccgcgcttggccgattcataatgcagct 5590
Db 7061 AAGAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTTCATTAATGCAGCT 7120
QY 5591 ggcaacgacaggtttcccgactggaaagcggcgagtgagcgcaacgcaaatcaatgtgagtt 5650
Db 7121 GGCACGACAGGTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 7180
QY 5651 agctcactcattaggcacccccagggctttacactttatgcttccggcgtcgtatgtgtg 5710
Db 7181 AGCTCACTCATTAGGCACCCCCAGGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 7240
|||||

QY 5711 gaattgtgagcgataaacaatttcacacaggaacaacagctatgaccatgattacgcca 5767
|||||
Db 7241 GAATTGTGAGCGGATAACAATTTACACACAGAAACAGCTATGACCATGATTACGCCA 7297

RESULT 7
AX114872/c
LOCUS AX114872 8062 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 32 from Patent WO0129208.
ACCESSION AX114872
VERSION AX114872.1 GI:14031814
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 8062)
AUTHORS Kuehn,R., von Melchener,H. and Altschmied,J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 32 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)
FEATURES
source Location/Qualifiers
1..8062
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pRK76"
BASE COUNT 1960 a 2036 c 2121 g 1945 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 8062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactgcccgtcgcttttacaaactgcgtgactgggaaaaacccctggcggttacccaac 3190
|||||
Db 8045 CGAATTCACTGGCGCTCGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAAC 7986

QY 3191 ttaatgccttcagacacatccccctttcgccagctggcgtaataagcgaagagcccgca 3250
|||||
Db 7985 TTAATGCGCTTGCGACACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCCGCA 7926

QY 3251 ccgatgcgccttcaccaacagttgcgagcctgaatggcggaatggcgctgatgcggtatt 3310
|||||
Db 7925 CCGATGCGCCCTTCCCAACAGTTGGCGAGCTGGAATGGCGAATGGCGCTGATGCGGTATT 7866

QY 3311 ttctccttacgcatctgtcggtatttcacaccgcataatggtgcactctcagtacaatct 3370
|||||
Db 7865 TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGTTGCACCTCTCAGTACAATCT 7806

QY 3371 gctctgatccgcatagttaagccagcccgacaccccgcaacaccccgctgacgcgccct 3430
|||||
Db 7805 GCTCTGATGCCGATAGTTAAGCCAGCCCCGACACCCGCCAACACCCCGTGACGCGCCCT 7746

QY 3431 gacggccttgctgctccccggcatccgcttacagacaagctgtgacgctctccgggagct 3490
|||||
Db 7745 GACGGGCTTGCTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 7686

QY 3491 gcatgtgtcagagggttttcaccgctcatcacccgaaacgcgcgagacgaaaggccctcgtga 3550
|||||
Db 7685 GCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCCGAGACGAAAGGGCCCTCGTGA 7626

QY 3551 tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggca 3610
|||||
Db 7625 TACGCCTATTTTATAGGTTAATGTCAATGATAATAATGGTTTCTTAGACGTCAGGTGGCA 7566

QY 3611 cttttcggggaaatgtgcgggaacccctatttgtttatttttctaaatacatctcaata 3670
|||||
Db 7565 CTTTTCGGGGAAATGTGCGGGAACCCCTATTGTATTATTTTCTAAATACATTCAAATA 7506

QY 3671 tgtatccgctcatgagacaataaacctgataaaatgcttcaataatattgaaaaaggaaga 3730
|||||
Db 7505 TGTATCCGCTCATGAGACAATAACCCCTGTAAATGCTTCAATAATATTGAAAAAGGAAGA 7446

QY 3731 gtatgagtattcaacatttcctggtgcgcccttattcccttttttcggcattttgccttc 3790
|||||
Db 7445 GTATGAGTATCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTCGGGCATTTTGCCTTC 7386

QY 3791 ctgtttttgctcaccagaaacgcgtggtgaaagtaaaagatgctgaagatcagttgggtg 3850
|||||
Db 7385 CTGTTTTTGTCAACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 7326

QY 3851 cacgagtggtttacatcgaactggatctcaacagcggtaagatcccttgagagttttcgcc 3910
|||||
Db 7325 CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCCTTGAGAGTTTTCGCC 7266

QY 3911 ccgaagaacgttttccaatgatgagcaccttttaaaagtctctgctatgtggcggtattat 3970
|||||
Db 7265 CCGAAGAACGTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 7206

QY 3971 cccgtattgacgcccgggcaagagcaactcgtgcgcgcacatacacactatttcagaatgact 4030
|||||
Db 7205 CCCGTATTGACGCGCGGCAAGAGCAACTCGTCGCCGCATACACTATTCTCAGAATGACT 7146

QY 4031 tggttgagtactcaccagtcacagaaaaagcaatctttacggatggcatgacagtaagagaat 4090
|||||
Db 7145 TGGTTGAGTACTCACAGTACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT 7086

QY 4091 tatgcagtgctgccataaacatgagtataaacactgcggccaaacttacttctgcacaacga 4150
|||||
Db 7085 TATGCAGTGTCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA 7026

QY 4151 tcggaggaccgaaggagctaacccgttttttcgcaacaatgggggatcatgtaaactcgcc 4210
|||||
Db 7025 TCGGAGGACCGAAGGAGCTAACCCGCTTTTTCACAAACATGGGGGATCATGTAACCTCGCC 6966

QY 4211 ttgatcgttggaacccggagctgaatgaagccataccaaacgacgagcgtgacaccacga 4270
Db 6965 TTGATCGTTGGAAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGA 6906

QY 4271 tgcctgtagaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctag 4330
Db 6905 TGCCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 6846

QY 4331 ctccccggcaacaattaatagactggatggaggcgggataaaagtgcaggaccacttctgc 4390
Db 6845 CTTCGCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACCTTCTGC 6786

QY 4391 gctcggcccttcggcgtggctgtttattgctgataaaactctggagccgggtgagcgtgggt 4450
Db 6785 GCTCGGCCCTTCGGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT 6726

QY 4451 ctcgcggtatcattgcagcactggggccagatggtaagccctcccgatcgtagttatct 4510
Db 6725 CTCGCGGTATCATTCAGCACACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6666

QY 4511 acacgacgggagtcaggcaactatggatgaacgaaatagacagatcgtcgatagagtg 4570
Db 6665 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGTGAGATAGGTG 6606

QY 4571 cctcactgattaaagcattggtaactgtcagaccaagtcttactcatatatacttttagattg 4630
Db 6605 CCTCACTGATTAAAGCATTTGGTAACCTGTACACCAAGTTTACTCATATATACTTTAGATTG 6546

QY 4631 atttaaaacttcattttttaatttaaaaggatcaggtgaagatcccttttgataaactc 4690
Db 6545 ATTTAAACATTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCA 6486

QY 4691 tgacaaaaatcccttaacgtgagtttctccactgagcgtcagaccccgtagaaaaa 4750
Db 6485 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCACCTGAGCGTCAGACCCCGTAGAAAAA 6426

QY 4751 tcaaaaggatcttcttgagatccctttttcttgcgcgtaaatctgctgcttgcaaaaaa 4810
Db 6425 TCAAAAGGATCTTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTGCAAAACAAAA 6366

Qy	4811	aaccacgcgtaccagcggtgttctgttgcgggatcaagagctaccaactcttttccga	4870
Db	6365	AACCACCGCTACCAGCGGTGGTTTGTTCGGGATCAAGAGCTACCAACTCTTTTCCGA	6306
Qy	4871	aggtaactggcttcagcagagcgcataccaaaatactgtccctctagtgtagccgtagt	4930
Db	6305	AGGTAAC TGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCCTCTAGTGTAGCCGTAGT	6246
Qy	4931	taggccaccacttcaagaactctgtagcaccgcctacatacactcgtctcgttaactcgt	4990
Db	6245	TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACTCTGCTCTGCTAATCCTGT	6186
Qy	4991	taccagtggctgctgccagtggcgataaagtctgtctcttaccgggttggaactcaagacgat	5050
Db	6185	TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTGGACTCAAGACGAT	6126
Qy	5051	agttaccggataaaggcgcagcggctcgggctgaacgggggttcgtgcacacagcccgct	5110
Db	6125	AGTTACCGGATAAAGCGCAGCGGTGGGCTGAACGGGGGTTCTGTGCACACAGCCCAGCT	6066
Qy	5111	tggagcgaacgacctaacaccgaactgagatacctaagcgtgagctatgagaaagcgcca	5170
Db	6065	TGGAGCGAAGACGACTACACCGAACTGAGATACCTACAGCGTGAGTATGAGAAAGCGCCA	6006
Qy	5171	cgcttccggaaggagaaaggcgcaggtatccggtaagcggcagggctcggaacaggag	5230
Db	6005	CGCTTCCGGAAGGGAGAAAGCGGCACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAG	5946
Qy	5231	agcgcagcagggagcgttccagggggaaacgcctgggtatcctttatagtcctgtcgggttc	5290
Db	5945	AGCGCACGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCTGTTCGGGTTTC	5886
Qy	5291	gccacctgacttgagctcgatttttgtgatgctcgtcagggggcgagcctatgga	5350
Db	5885	GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA	5826
Qy	5351	aaaacgcagcaacgcggcctttttacgggttcctggccttttgcggccttttgctcaca	5410
Db	5825	AAAACGCAGCAACGCGGCCTTTTACGGTTCCTGGCCCTTTTGTGGCCTTTTGCTCACA	5766
Qy	5411	tgttcttctcgtgcttatccccctgattctgtggataaacctgattaccgcctttgagtga	5470
Db	5765	TGTTCTTTCTTCTGCGTTATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG	5706
Qy	5471	ctgataccgctcgcgcgcagccgaacgcagcagcgcagcagtcagtcagtcagcagcgg	5530
Db	5705	CTGATACCGCTCGCCGCAGCCGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGG	5646
Qy	5531	aagagcgcccaataacgcaaaaccgcctctccccgcgcgttggccgatcattcaataatgcagct	5590
Db	5645	AAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCTTGCGCCGATTCAATATGCAGCT	5586
Qy	5591	ggcacgcagggtttccccgactggaagcgggcagtcagtcagtcagcgaacgaattaatgtgagt	5650
Db	5585	GGCACGACAGGTTTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT	5526
Qy	5651	agctcactcattagggcacccccaggctttacacctttatgcttccggctcgtatgttgtg	5710
Db	5525	AGCTCACTCATTAGGCACCCCAAGGCTTTACACTTTATGCTTCCGSGCTCGTATGTTGTGTG	5466
Qy	5711	gaattgtgagcggataaacaatttcacacagggaacagctatgacctgattacgcca	5767
Db	5465	GAATTGTGACGGGATAACAATTTACACACAGGNAACAGCTATGACCATGATTACGCCA	5409
RESULT 8			
AX114871/c			
LOCUS AX114871 8153 bp DNA linear PAT 11-MAY-2001			
DEFINITION Sequence 31 from Patent WO0129208.			
ACCESSION AX114871			
VERSION AX114871.1 GI:14031813			
KEYWORDS .			
SOURCE synthetic construct.			

ORGANISM	synthetic construct		
REFERENCE	artificial sequence.		
AUTHORS	1 (bases 1 to 8153)		
TITLE	Kuehn,R., von Melchener,H. and Altschmied,J.		
JOURNAL	Conditional gene trapping construct for the disruption of genes		
FEATURES	Patent: WO 0129208-A 31 26-APR-2001;		
source	ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)		
	Location/Qualifiers		
	1..8153		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="vector pRK74"		
BASE COUNT	1976 a	2133 c	2067 g
ORIGIN	1977 t		
Query Match 45.7%; Score 2637; DB 6; Length 8153;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	3131	cgaattcactggccgctcgtttttacaacgtcgtgactgggaaaaacccctggcgttacccaac	3190
Db	8136	CGAATTCACTGGCCGTCGTTTACAAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAAC	8077
Qy	3191	ttaatgccttgcagcacatcccccttttcgacagctggcgtaatagcgaagagcccgca	3250
Db	8076	TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA	8017
Qy	3251	ccgatacgcccttcccaacagttgcgcagcctgaatggcgaaatggcgctgatcggtatt	3310
Db	8016	CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTGATCGGTATT	7957
Qy	3311	ttctccttacgcatctgtgcggtatttcacacccgcataatggtgcactctcagtacaatct	3370
Db	7956	TTCTCCTTACGCATCTGTGCGGTATTTACACCGCATATGGTGCACCTCTCAGTACAATCT	7897
Qy	3371	gctctgatgcgcgcatagttaagccagcccgacacccgcgaacaccccgctgacgcgccct	3430
Db	7896	GCTCTGATGCGGCATAGTTAAGCCAGCCCCCGACACCCGCCAACCCGCTGACGCGCCT	7837
Qy	3431	gacgggcttgctcgtctcccggcatccgcttacagacaagctgtgaccgctctccgggagct	3490
Db	7836	GACGGGCTTGCTGCTCCCCGCATCCGCTTACAGACAAAGCTGTGACCGTCTCCGGGAGCT	7777
Qy	3491	gcatgtgtcagaggtttttcacccgtcatcacccgaacacgcgcgagacgaaggccctcgtga	3550
Db	7776	GCATGTGTCAAGAGGTTTTCACCGTCTACCCGAAACGCGGAGACGAAAGGGCCTCGTGA	7717
Qy	3551	tacgcctattttataggttaatgtcatgataataatggtttcttagcagtcaggtggca	3610
Db	7716	TACGCCATTATTTATAGGTTAATGTCAATGATAAATAATGGTTTCTTAGACGTCAGGTGGCA	7657
Qy	3611	cttttcgggaaaatgtcgcggaaacccctatttgttatttttctaatacatcacaata	3670
Db	7656	CTTTTCGGGAAAATGTGCGGGAACCCCTATTGTGTTATTTCCTAAATACATTCAAATA	7597
Qy	3671	tgtatccgctcatgagacaataaaccttgataaaatgcttcaataatatgaaaaaggaaga	3730
Db	7596	TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATGAAAAAGGAAGA	7537
Qy	3731	gtagagtattcaacatttcctcgtgtcgcccttattcccttttttggcgcattttgccttc	3790
Db	7536	GTATGAGTATCAACATTTCCGTTGTCGCCCTATTCCCTTTTTCGGGCATTTTGCCTTC	7477
Qy	3791	ctgtttttgtcaccacagaaacgctggtgaaaagtaaaagatcgtgaagatcagttgggtg	3850
Db	7476	CTGTTTTGTCTCACCACAGAAACGCTGGTGAAAAGTAAAAAGATGCTGAAGATCAGTTGGGTG	7417
Qy	3851	cacgagtggttacatcgaaactggatcctcaacagcggtaagatccttgagagtttccgcc	3910
Db	7416	CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC	7357
Qy	3911	ccgaagaacgttttcccaatgatgagcacttttaaaagtctgctatgtggcgcggtattat	3970

Db	7356	CGAAGAACGTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGCGGTATTAT	7299
QY	3971	cccgtattgacgcgcgggcaagagcaactcggctgcgcgcatacacactattctcagaatgact	4030
Db	7296	CCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACT	7237
QY	4031	tggttgagtactcacaggtcacagaaaaagcatctttacggatggcatgacagtaagagaat	4090
Db	7236	TGGTTGAGTACTCACCAGTACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT	7177
QY	4091	tatgcagtgctgccataaccatgagtataaacactcggcgccaacttacttctgcacaacga	4150
Db	7176	TATGCAGTGCTGCCATAACCATGAGTGATAAACACTGCGGCCAACTTACTTCTGACAACGA	7117
QY	4151	tcggaggaccgaaggactaacccgcttttttgcacaacatcggggatcatgtaactcgcc	4210
Db	7116	TCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCC	7057
QY	4211	ttgatcgttgggaacccggagctgaatgaagccataccaaacgacgagcgtgacaccacga	4270
Db	7056	TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA	6997
QY	4271	tgcctgtagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctag	4330
Db	6996	TGCTGTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG	6937
QY	4331	cttcccgccgcaacaataatagactggatggaggcggaataaagttgcaggaccactctctgc	4390
Db	6936	CTTCCCGGCAACAATAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCCTTCTGTC	6877
QY	4391	gctcggccctccggctggctggtttattgctgataaaactctggagccggtgagcgtgggt	4450
Db	6876	GCTCGGCCCTTCCGGCTTGGCTGGTTATTGCTGTATAAATCTGGAGCCGCTGAGCGTGGGT	6817
QY	4451	ctcgcggtatcattgcagcactggggccagatggttaagccctcccgtatcgtagttatct	4510
Db	6816	CTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTTAAGCCCTCCCGTATCGTAGTTATCT	6757
QY	4511	acacgacggggagtcaggcaactatggatgaacgaaaaatagacagatcgctgagataggtg	4570
Db	6756	ACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG	6697
QY	4571	cctcactgattaagcatbggtaactgtcagaccagtttactcatatatacttttagattg	4630
Db	6696	CCTCACTGATTAAAGCAFTGGTAACCTGTCAAGCAAGTTTACTCATATATACTTTAGATTG	6637
QY	4631	atttaaaacttcattttaatttaaaaggatctagggtgaagatccctttttgataaatctca	4690
Db	6636	ATTTAAAACTTCATTTTAAATTTAAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCA	6577
QY	4691	tgaccaaaaatcccttaacgtgagttttctcgttccactgagcgtcagacccccctagaaaaa	4750
Db	6576	TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA	6517
QY	4751	tcaaaggatcttcttgagatcccttttttctgcgcgtaaatctgctgcttgcaaacaaaaa	4810
Db	6516	TCAAAGGATCTTCTTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCACAAACAAA	6457
QY	4811	aaccaccgctaccagcgggtgtttgtttgcccggatcaagagctaccaactctttttccga	4870
Db	6456	AACCACCGCTACCAGCGGTGTTGTTTGGCCGGATCAAGAGCTACCAACTCTTTTTCCTCGA	6397
QY	4871	aggtaactggcttcagcagcgcagataccaaaatactgtccttctagttagccgtagt	4930
Db	6396	AGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTTAGTGTAGCCGTAGT	6337
QY	4931	taggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgtctaactctgt	4990
Db	6336	TAGGCCACCACCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAAATCCTGT	6277
QY	4991	taccagtggtgctgcccagtggcgataagtcgttcttaccgggttggactcaagacgat	5050

[illegible]

9	RESULT
9	XXP35SGFP
9	LOCUS
9	DEFINITION
9	ACCESSION
9	VERSION
9	KEYWORDS
9	SOURCE
9	ORGANISM
9	REFERENCE
9	AUTHORS
9	TITLE
9	JOURNAL
9	REFERENCE
9	AUTHORS
9	TITLE
9	JOURNAL
9	REFERENCE
9	AUTHORS
9	TITLE

JOURNAL Submitted (05-JUN-1995) Paul A. Kitts, CLONTECH Laboratories, Inc., 4030 Fabian Way, Palo Alto, CA 94303, USA
COMMENT This vector can be obtained from CLONTECH Laboratories, Inc., 4030 Fabian Way, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424-8222 or (800) 662-2566, extension 3. This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.

FEATURES Location/Qualifiers
source 1..4518
/organism="Cloning vector p35S-GFP"
/plasmid="p35S-GFP"
/db_xref="taxon:41037"
896..1612
/gene="gfp"
896..1612
/gene="gfp"
/codon_start=1
/transl_table=11
/product="green fluorescent protein"
/protein_id="AA69675.1"
/db_xref="GI:894140"
/translation="MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATYVKLTLLKFICTGKLPVPWPTLVTFSYGVQCFSRYPDRHKQHDFFKSAWPEGVVQERTIFFKDDGNYKTRAEVKFEGDILVNRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQNTPIGDGPPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK"
2485..3345
/gene="bla"
2485..3345
/gene="bla"
/function="confers resistance to ampicillin"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA69676.1"
/db_xref="GI:894141"
/translation="MSIQHFRVALIPFFAAFLCPVFAHPETLVKVKDAEDQLGARVGYIELDLSNGKILLESFRRPEERFPMMSTFKVLLCGAVLSRIDAGQQQLGRRIRHYSONDLVYESPVTEKHLTDGMFVRELCSAAITMSDNTAANLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTPMPVAMATTLRKLLTGGELLTLASRQQLIDWMEADKVVAGPLLRSLPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSOATMDERNRQIAEIGASLIKHW"
BASE COUNT 1287 a 1055 c 1042 g 1134 t

Query Match 45.7%; Score 2635; DB 12; Length 4518;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3132 gaattcactggccgctgcttttacaaactgctgactgggaaacccctggcgttacccaact 3191
Db 1884 GAATTCACTGGCCGTCGTTTACAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAACT 1943
QY 3192 taatgccttgagcacatcccccttcgcagctggcgtaataagcgaagagcccgac 3251
Db 1944 TAATCGCCTTCCCAACAGTTGGCAGCGCTGAATGGCGAATAGCGAAGAGCCCGCAC 2003
QY 3252 cgatgccttcccaacagttgcgagcctgaatggcgaatggcgctgatgcggtattt 3311
Db 2004 CGATCGCCTTCCCAACAGTTGGCAGCGCTGAATGGCGAATGGCGCTGATCGGTATTT 2063
QY 3312 tctccttaagcatctgtgcggtatttcacacccgcataatggtgcactctcagtacaatctg 3371
Db 2064 TCTCCTTACGCATCTGTGCGGTATTTACACCGCATATGGTGCACCTCTCAGTACAATCTG 2123
QY 3372 ctctgatgcgcgatagttaagccagcccgacacccgcccaacaccgcgtgacgcgccctg 3431

Db 2124 CTCTGATGCCGCATAGTTAAGCCAGCCCGACACCCGCAACACCCGCTGACGCGCCTG 2183
QY 3432 acgggcttctgctcccgcatccgcttacagacaagctgtgacgctcctccggagctg 3491
Db 2184 ACGGGCTTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTG 2243
QY 3492 catgtgcagaggtttcacccgtcatcacgaaacgcgcgagacgaaggccctcgtgat 3551
Db 2244 CATGTGTCAGAGGTTTTCACCGTTCATACCGAAACGCGGAGACGAAAGGCCCTCGTGAT 2303
QY 3552 acgcctattttataggttaatgtcatgataataatggtttcttagacgctcaggtggcac 3611
Db 2304 ACGCCTATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCAC 2363
QY 3612 ttttcggggaatgtgcgcggaacccctatttggtttttttcttaataacattcaaatat 3671
Db 2364 TTTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTCAATAT 2423
QY 3672 gtatccgctcatgagacaataaacctgataaaatgcttcaataatatgaaaaaggaagag 3731
Db 2424 GTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAG 2483
QY 3732 tatgagtattcaacatttccgtgtgcgcccttattcccttttttgcggcattttgccttcc 3791
Db 2484 TATGAGTATTCAACATTTCGCTGTGCGCCTATTCCCTTTTTCGSGCATTTTGCCTTCC 2543
QY 3792 tgtttttgctcaccagaaacgcgtggtgaaagtaaaagatgctgaagatcaggttggtgc 3851
Db 2544 TGTTTTGTGTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGC 2603
QY 3852 acgagtgggttacatcgaaactggatctcaacacgcggtaagatcccttgagagttttcgccc 3911
Db 2604 ACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC 2663
QY 3912 cgaagaacgttttccaatgatgagcacttttaaagtctgctatgtggcgcggtattatc 3971
Db 2664 CGAAGAACGTTTCCAAATGATGAGCACTTTTAAAGTTCTGTATGTGGCGCGGTATTATC 2723
QY 3972 ccgtattgacgcccgggcaagagcaactcgggtgcgcgcatacactattctcagaatgactt 4031
Db 2724 CCGTATTGACGCGGGCAAGAGCAACTCGTGCGCGCATACACTATTCTCAGAANTGACTT 2783
QY 4032 ggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaatt 4091
Db 2784 GGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATT 2843
QY 4092 atgcagtgtgccataaacatgagtataaacactgcggccaaacttacttctgacaacgat 4151
Db 2844 ATGCAGTGTGCCATAACCATGAGTGATGAACACTGCGGCCCACTTACTTCTGACAACGAT 2903
QY 4152 cggaggaccgaaggagtaaccgctttttgcacaacatgggggatacatgtaactcgcct 4211
Db 2904 CGGAGGACCGAAGGAGCTAACCGCTTTTTCACAACATGGGGGATCATGTAACTCGCCT 2963
QY 4212 tgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgat 4271
Db 2964 TGATCGTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGAT 3023
QY 4272 gcctgtagaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttacttagc 4331
Db 3024 GCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTTAGC 3083
QY 4332 ttcccgccctcccgctggctggctggatggagcggaataaaagtgcagagaccacttctgcg 4391
Db 3084 TTCCCGGCAACAATTAATAGACTGGATGGAGCGGGGATAAAGTTGCAGGACCACCTTCTGCG 3143
QY 4392 ctcggcccttcccgctggctggctgttttatgctgataaaatctggagccggtgagcgtgggtc 4451
Db 3144 CTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCCCGGTGAGCGTGGGTC 3203
QY 4452 tcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatcta 4511

Db	2125	CTCTGATGCCGATAGTTAAGCCAGCCCGACACCCGCCAACACCCCGTGACGGCCCTG	2184
QY	3432	acgggctgtctgtctcccgcatccgcttacagacaagctgtgacctctccggagctg	3491
Db	2185	ACGGGCTTGTCTGCCGTCACGCTACAGACAAGCTGTACCGCTCCGGGAGCTG	2244
QY	3492	catgtgtcagaggttttccaccgtcatcacggaacgcgcgagacgaaagggcctcgtgat	3551
Db	2245	CATGTGTcAGAGGTTTTCACCGTCATCCGGAACCCCTATTGTTTCTTAGACGTCAGGTGGCAC	2304
QY	3552	acgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcac	3611
Db	2305	ACGCCCTATTTTATAGGTTAATGTcATGATAATAATGGTTTCTTAGACGTCAGGTGGCAC	2364
QY	3612	ttttcgggaaatgtgcgcggaacccctatttgtttatttttctaaatacatat	3671
Db	2365	TTTTTCGGGAAATGTGCGCGAACCCCTATTGTTTATTTTCTAAATACATTCAATAT	2424
QY	3672	gtaaccgtcatgagacaataaacccctgataaatagtcttcaataataatgaaaggaagag	3731
Db	2425	GTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAG	2484
QY	3732	tatgagtattcaacatttcctgtgtcgccctatttcccttttttgccgcaatttgccttc	3791
Db	2485	TATGAGTATTCAACATTTCCGTGTGCGCCCTATTTCCTTTTTTGGGCATTTTGCCTTCC	2544
QY	3792	tgtttttgtcaccagaaacgctggtgaaagtataagatgctgaagatcagttgggtgc	3851
Db	2545	TGTTTTTGTcACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGC	2604
QY	3852	acgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagtttgcgcc	3911
Db	2605	ACGAGTGGSTTACATCGAACTGGATCTCAACAGCGGTAAAGTCCTTGAGAGTTTTCGCC	2664
QY	3912	cgaagaacgttttccaatgatgagcaccttttaaagttctgtctatgtggcgcgtattatc	3971
Db	2665	CGAAGAACGTTTTCCAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGCGGGTATTATC	2724
QY	3972	ccgtattgacgcgcgggcaagagcaactcggctgcgcgcatacacactattctcagaatgactt	4031
Db	2725	CCGTATTGACGCCGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAAATGACTT	2784
QY	4032	ggttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaatt	4091
Db	2785	GGTTGAGTACTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATT	2844
QY	4092	atgcagtgtgccataaacatgagtgaataacactcggcgaacttactctgacaaacgat	4151
Db	2845	ATGCAGTGTGCCATAAACCATGAGTGATAACACTCGGCCCAACTTACTTCTGACAAACGAT	2904
QY	4152	cggaggaccgaagagctaaccgccttttttgcaacaacatgggggatcatgtaactcgcct	4211
Db	2905	CGGAGGACCGAAGGAGCTAACCCGCTTTTTCACACAACATGGGGGATCATGTAACTCGCCT	2964
QY	4212	tgatcgttggaaccggagctgaatgaagccataaccaaaacgacgagcgtgacacccacgat	4271
Db	2965	TGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCCACGAT	3024
QY	4272	gcctgtageaatggcaacaacgttgcgcaactattaactggcgaactacttactctagc	4331
Db	3025	GCCTGTAGCAATGGCAACAACGTTTGCGCAAACTATTAACTGGCGNACTACTTACTCTAGC	3084
QY	4332	ttcccggaacaataatagactggatggagggcggataaaagttgcaggaccacttctgcg	4391
Db	3085	TTCCCGGCACAATAATAGACTGGATGGAGCGCGGATAAAGTTGCAGGACCACCTCTGCG	3144
QY	4392	ctcggcccttcggctggctggtttattgtctgataaaatctggagccggtgagcgtgggtc	4451
Db	3145	CTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCCGGTGAGCGTGGGTC	3204
QY	4452	tcgcgggtatcatgcagcacactggggccagatgggtaagccctcccgatatcgtagttatcta	4511

Db	3205	TCGCGGTATCATTCAGCACTGGGGCCAGATGCTAAGCCCCCTCCCGTATCGTAGTTATCTA	3264
QY	4512	cacgacggggagtcaggcaactatggatgaacgaaaatagacagatcgctgagataggtgc	4571
Db	3265	CACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGC	3324
QY	4572	ctcactgattaaacatttgtaactgtcagaccaaagtttactcatatatacttttagattga	4631
Db	3325	CTCACTGATTAAAGCATTGGTAACGTGTcAGACCAAGTTTACTCATATATACTTTTAGATTGA	3384
QY	4632	tttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttttgataatctcat	4691
Db	3385	TTTAAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCAT	3444
QY	4692	gacaaaaatccccttaacgtgagtttctgttccactgagcgtcagacccccgtagaaaaagat	4751
Db	3445	GACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGAT	3504
QY	4752	caaaggatcttcttgagatcccttttttctgcgcgtaaatctgctgcttgcaaaaaaaa	4811
Db	3505	CAAAGGATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGCACAAACAAAAA	3564
QY	4812	accaccgctaccagcgggtggtttgtttgcggatcaagagctaccaactctttttccgaa	4871
Db	3565	ACCACCGCTACCACGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAA	3624
QY	4872	ggtaaactggcttcagcagagcgcagataaccaataactgtccttctagtgtagccgtagtt	4931
Db	3625	GGTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCGGTAGTT	3684
QY	4932	aggccaccacttcaagaaactctgtagcacccgctacatacctcgcctcgtcgtcgtcgtgtt	4991
Db	3685	AGGCCACCACCTCAAGAACTCTGTAGCACCCGCTACATACTCGCTCTGCTAATCCTGTT	3744
QY	4992	accagtggctgctgcagtggcgataagtcgtcttaccgggttggaactcaagacgata	5051
Db	3745	ACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTGGACTCAAGACGATA	3804
QY	5052	gttaccgggataaaggcgcagcggtcgggctgaacggggggttcgtgcacacacagcccagctt	5111
Db	3805	GTTACCGGATAAGGCGCAGCGGTTCGGCTGAACGGGGGGTTCGTGCACACAGCCACGCTT	3864
QY	5112	ggagcgaacgacctacaccgaactgagataacctacacgctgagctatgagaaagcgcac	5171
Db	3865	GGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCAC	3924
QY	5172	gcttcccgaaagggaagggcacaggtataccggtaagcggcagggtcgggaacaggaga	5231
Db	3925	GCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTTCGGAACAGGAGA	3984
QY	5232	gcgcacgaggaggcttccagggggaaacgcctggtatcttatagtcctgtcggggttcg	5291
Db	3985	GCGCACGAGGGAGCTTCCAGGGGGAACGCCCTGSTATCTTTATAGTCCCTGTGCGGTTTCG	4044
QY	5292	ccacctctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatggaa	5351
Db	4045	CCACCTCTGACTTGACGCTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCCCTATGGAA	4104
QY	5352	aaacgccagcaacgcggccttttaacggttccctggcccttttgccttttgcacacat	5411
Db	4105	AAACGCCAGCAACGGCGCCTTTTACGGTTCCCTGGCCTTTTGTGGCCTTTTGTCTCACAT	4164
QY	5412	gttcttctcgtcgttataccccctgattctctgtggaataaccgtattaccgcctttgagtgagc	5471
Db	4165	GTTCTTCTGCGTTATCCCCCTGATTCTGTGGATAACCCGTATTACCGCCCTTTGAGTGAGC	4224
QY	5472	tgataccgctcgcgcagccgaacgacccgagcgcagcagtcagtgagcgcgagcgcgga	5531
Db	4225	TGATACCGCTCGCCGACGCCGAACGACCGGAGCGGAGTCAGTGAGCGAGGAAGCGGA	4284
QY	5532	agagcgcccaataacgcaaaacccgcctctccccgcgcggttgccgattcattaatgcagctg	5591
Db	4285	AGAGCGCCCAATACGCAAAACCCGCTCTCTCCCCCGCGGCTTGCCCGATTTCATTAAATCGCAGCTG	4344

||||| 3328 CGCGGTATCATTCAGCACTGGGGCCAGATGCTAAGCCCTCCCGTATCGTAGTTATCTAC 3387

Qy 4513 acgacggggagtcaggcaactatggatgaacgaaaaatagacagatcgctgagatagtgcc 4572

Db 3388 ACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC 3447

Qy 4573 tcaactattaagcattgtaactgcagaccaagtttactcataatataacttttagattgat 4632

Db 3448 TCACTGATTAAGCATTTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT 3507

Qy 4633 ttaaaacttcatttttaatttaaaaggatctaggtgaagatccctttttgataatctcatg 4692

Db 3508 TTAATAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATG 3567

Qy 4693 accaaaaacccttaacgtgagtttttcgctccactgagcgtcagacccccgtagaaaaagatc 4752

Db 3568 ACCAAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATC 3627

Qy 4753 aaaggatcttcttgagatcccttttttctgcgcgtaaatctgctgcttgcaaaacaaaaaa 4812

Db 3628 AAAGGATCTTCTTGAGATCCCTTTTTTCTGCGCGTAATCTGCTGTGCAACAAAAAAA 3687

Qy 4813 ccaccgctaccagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4872

Db 3688 CCACCGCTACCAGCGGTGTTGTTGTCGCGGATCAAGAGCTACCAACTCTTTTCCGAAG 3747

Qy 4873 gtaactggcttcagcagagcgcagataccaaaaatactgtccttctagtgtagccgtagtta 4932

Db 3748 GTAACGGCTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTA 3807

Qy 4933 ggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctgtaacctgttta 4992

Db 3808 GGCCACCCTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCTGCTAATCCTGTTA 3867

Qy 4993 ccagtggctgctgccagtgccgataaagtgcgtgtcttaccgggttggaactcaagacgatatg 5052

Db 3868 CCACTGGTCTGCTGCCAGTGGCGATAAGTCTGTCTTTACCGGTTGGACTCAAGACGATAG 3927

Qy 5053 ttaccggataaaggcgcagcggctgcggctgaacgggggttcgtgcacacagccccagcttg 5112

Db 3928 TTACCGGATAAGGGCGCAGCGGTTCGGGCTGAACGGGGGTTTCGTGCACACAGCCCAGCTTG 3987

Qy 5113 gagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcggcacg 5172

Db 3988 GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCGCCACG 4047

Qy 5173 ctcccggaaggagaaaggcgggacaggtatccggttaagcggcgaggtgcggaacaggagag 5232

Db 4048 CTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTGCGGAACAGGAGAG 4107

Qy 5233 cgcacgaggagcttccagggggaaacgcctgggtatctttatagtctgctgctggttctgc 5292

Db 4108 CGCACGAGGAGACTTCCAGGGGGAACCCCTGGTATCTTTATAGTCTCTGTCGGGTTTCGC 4167

Qy 5293 cacctctgacttgagcgtcgatctttttgtgatgctcgtcaggggggggagcctatggaaa 5352

Db 4168 CACCTCTGACTTGAGCGTCGATTTTTTGTGATGCTCGTCAGGGGGGGGAGCCTATGGAAA 4227

Qy 5353 aacgccagcaacgcggcctttttacgggttcctggccttttgcgtggccttttgcctcacatg 5412

Db 4228 AACGCCAGCAACGGCGCCCTTTTACGGTTCTCTGGCCCTTTTGTCTGGCCCTTTGCTCACATG 4287

Qy 5413 ttctttcctgcgttatccccctgattctgtggataaccgattaccgcctttgagtgagct 5472

Db 4288 TTCTTTCTCGGTATCCCCCTGATTCGTGGGATAACCGTATTACCGCCTTTGAGTGAGCT 4347

Qy 5473 gataccgctgcgcgagccgaacgacgagcgcagcgcagtcagtcagtcagtcagtcagtcag 5532

Db 4348 GATACCGCTCGCGCAGCCGAACGACCGGCGCAGCGAGTCAGTGAGCGAGGAGCGGAA 4407

Qy 5533 gagcgcccaataacgcaaacccgcctctccccgcgcgttgcccgattcaattaatgcagctgg 5592

|||||

Db 4408 GAGCGCCCAATACGAAACCGCCTCTCCCGCGCGTTGGCCGATTTCATTAATGCAGCTGG 4467

Qy 5593 cagcacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttag 5652

Db 4468 CACGACAGGTTTCCCGACTGGAAAGCGGSCAGTGAGCGCAACGCAATTAATGTGAGTTAG 4527

Qy 5653 ctcaactcattaggcacccccaggcctttacactttatgcttccggcctgctatgtgtgtgga 5712

Db 4528 CTCACCTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGA 4587

Qy 5713 attgtgagcgggataaacaatttcacacaggaacacagctatgacctattacgcca 5767

Db 4588 ATGTGAGCGGATAACAATTTACACACAGGAAACAGCTATGACCATGATTACGCCA 4642

RESULT 12

E49323/c

LOCUS

DEFINITION E49323 3796 bp DNA linear PAT 31-JAN-2002

Infectious cDNA clone of North American porcine reproductive and respiratory syndrome (PRRS) virus and use thereof.

ACCESSION E49323

VERSION E49323.1 GI:18628054

KEYWORDS JP 2000I89178-A/32.

SOURCE unidentified plasmid.

ORGANISM unidentified plasmid

plasmids.

REFERENCE 1 (bases 1 to 3796)

AUTHORS Calvert,J.G., George,M. and Welshu,S.H.

TITLE Infectious cDNA clone of North American porcine reproductive and respiratory syndrome (PRRS) virus and use thereof

JOURNAL Patent: JP 2000I89178-A 32 11-JUL-2000;

PFIZER PROD INC

COMMENT

OS Plasmid

PN JP 2000I89178-A/32

PD 11-JUL-2000

PF 21-DEC-1999 JP 1999362186

PR 22-DEC-1998 US 60/113345

PI J GUREGORI CALVERT,MICHAEL GEORGE,SHAOKUN HWANG WELSHU PC

C12N15/09,A61K39/12,A61K48/00,A61P31/12,C12N1/15,C12N1/19, PC

C12N1/21,

PC C12N5/10,C12N15/00,C12N5/00

CC

FH Key Location/Qualifiers

FT source 1..3796

FT /organism='Plasmid'.

FEATURES

source

1..3796

/organism="unidentified plasmid"

/db_xref="taxon:45202"

BASE COUNT 954 a 932 c 954 g 956 t

ORIGIN

Query Match 45.6%; Score 2630; DB 6; Length 3796;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3138 actgcccgtcgtttttacaacgctgctgactgggaaaaaccctgctgcttaccacacttaacgc 3197

Db 3796 ACTGGCCGTCGTTTACAACGTCGTGACTGGGAAAAACCCTGGCGTTACCCAACTTAATCG 3737

Qy 3198 ccttgacgacatcccccttttcgccagctgctgtaataagcgaagagccgcacccgacgc 3257

Db 3736 CCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCG 3677

Qy 3258 cccttcccaacagttgcgcagcctgaatggcgaatggcgcctgatgcggtattttctcct 3317

Db 3676 CCCTTCCCAACAGTTGCGCAGCCCTGAATGGCGAATGGCGCTGATCGCGGTATTTCTCCT 3617

Qy 3318 tacgcattctgctgggtattttcacacccgcataatggtgcactctcagtaacaatctgctctga 3377

Db 3616 TAGGCATCTGTGGGTATTTACACACCGCATATGTTGGTGCACCTCTCAGTACAAATCTGCTCTGA 3557

QY 3378 tgccgcatagtttaagccagccccgcagacaccgcgccaacacccccgctgacgcgccttgacgggc 3437
Db 3556 TGCCCGCATAGTTAAGCCAGCCCCCGACACCCGCCAACACCCGCTGACGGCCCTGACGGGC 3497
QY 3438 ttgtctgctcccgccatccgccttacagacaagctgtgaccgtctccgggagctgcatgtg 3497
Db 3496 TTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTG 3437
QY 3498 tcagaggttttcaccgctcatcacccgaacgcgcgagacgaagggcctcgtgatacgcct 3557
Db 3436 TCAGAGGTTTTCACCGTTCATCACCGAAACGCGCGAGACGAAAGGGCCCTCGTGATACGCCT 3377
QY 3558 attttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcg 3617
Db 3376 ATTTTATAGGTTAATGTATGATAATAATGGTTTCTTAGACGTCAGTGGCACTTTTCG 3317
QY 3618 gggaaatgtgcgcggaacccctatttgtttatttttctaaatatattcaaatatgtatcc 3677
Db 3316 GGGAAATGTGCGGGAACCCCTATTGTGTATTTTTCTAAATACATTCAAAATATGTATCC 3257
QY 3678 gctcatgagacaataaacctgatataatgcttcaataataattgaaaaaggaagatagag 3737
Db 3256 GCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATGAG 3197
QY 3738 tattcaacatttcgcgtgtgcgcccttatcccttttttgcggcattttgccttcctgtttt 3797
Db 3196 TATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTGTGGGCATTTTGCCTTCTCTGTTT 3137
QY 3798 tgctcaccagaacacgctggtgaaagttaaagatgctgaagatcagttgggtgcacgagt 3857
Db 3136 TGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT 3077
QY 3858 gggttacatcgaaactggatctcaacagcgggtaagatcccttgagagttttcgcgccgaaga 3917
Db 3076 GGGTTACATTCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCGGAAGA 3017
QY 3918 acgttttccaatgatgagcaccttttaagttctgctatgtggcgcggtattatcccgat 3977
Db 3016 ACGTTTTCGAATGATGAGCACITTTTAAAGTTCTGCTATGTGGCGCGTATTATCCCGTAT 2957
QY 3978 tgacgcggggaagagcaactcggctgcgcgcatacacactattctcagaatgacttggttga 4037
Db 2956 TGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTGA 2897
QY 4038 gtactcaccagtcacagaaaaagcatcttaacggatggcatgacagtaagagaattatgcag 4097
Db 2896 GTACTCACCACTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAG 2837
QY 4098 tgctgccataaccatgagtataaacactgcggccaaacttacttctgcacaacgatcggagg 4157
Db 2836 TGCTGCCATAACCATGAGTGATAAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGG 2777
QY 4158 accgaaggagctaaccgcttttttgacaacaatggggatcatgtaactcgcttgatcg 4217
Db 2776 ACCGAAGGAGCTAAACCGCTTTTTCACAACATGGGGATCATGTAACTCGCCTTGATCG 2717
QY 4218 ttgggaaccggagctgaatgaagccataaccaaacgacgagcgtgacaccacgatgcctgt 4277
Db 2716 TTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGATGCCCTGT 2657
QY 4278 agcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctagcttcccg 4337
Db 2656 AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGGAACACTACTACTCTAGCTTCCCG 2597
QY 4338 gcaacaattaatagactggatggagcgggataaaagtgcaggacacacttctgcgtcggc 4397
Db 2596 GCAACAATTAATAGACTGGATGGAGCGCGGATAAAAGTTGCAGGACCACTTCTGCGCTCGGC 2537
QY 4398 ccttcggctggctggtttattgctataaaatctggagccggtgagcgtgggtctcgcgg 4457
Db 2536 CCTCCGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCCCGGTGAGCGTGGGTCTCGCGG 2477
QY 4458 tatcatgcagcactggggccagatgggtaagccctcccgctatcgtaggttatctacacgac 4517

Db 2476 TATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC 2417
QY 4518 ggggagtcaggcaactatggatgaacgaataagacagatcgctgagataggtgcctcact 4577
Db 2416 GGGGAGTCAGSCAACTATGGATGAACGAATAGACAGATCGCTGAGTAGGTGCCTCACT 2357
QY 4578 gattaagcattgggtaactgtcagaccaagtttactcatatatatacttttagattgatttaa 4637
Db 2356 GATTAAGCATTTGGTAACCTGTACAGACCAAGTTACTCATATATATCTTTAGATTGATTTAA 2297
QY 4638 acttcaatttttaatttaaaaggatctaggtgaagatcccttttbtgataaatctcatgaccaa 4697
Db 2296 ACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATGACCAA 2237
QY 4698 aatcccttaacgtgagtttctgtccactgagcgtcagacccccgtagaaaaagatcaaaagg 4757
Db 2236 AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATCAAAAGG 2177
QY 4758 atcttcttgagatcccttttttctgcgcgtaactctgctgtgcacacaaaaaaaccacc 4817
Db 2176 ATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACCCACC 2117
QY 4818 gctaccagcggtggttgttgcgcggatcaagagctaccaactcttttccgaaggtaac 4877
Db 2116 GCTACCAGCGGTGTTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC 2057
QY 4878 tggcttcagcagagcgcagatataccaaaatactgtcccttcttagtgcgtagcttagtgcca 4937
Db 2056 TGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCA 1997
QY 4938 ccacttcaagaaactctgttagcaccgcctacaataacctgcctcgtcgtctaactcgttaccagt 4997
Db 1996 CCACTTCAAGAACTCTGTAGCACCCGCCCTACATACTCTGCTCTGTCTGTATCTCTTACCAGT 1937
QY 4998 ggctgctgccagtgccgataagtcgtgtcttaccgggttggaactcaagacgatagttacc 5057
Db 1936 GGCTGCTGCCAGTGGCGATAAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAGTTACC 1877
QY 5058 ggataaggcgcagcggctcgggctgaacggggggttcgtgcacacagccacgcttgagcg 5117
Db 1876 GGATAAGGCGCAGCGGTGCGGCTGAACGGGGGTTGCTGCACACAGCCACGCTTGAGCG 1817
QY 5118 aacgacctacacgaactgagatataacctacagcgtgagctatgagaaagcgcacgcttcc 5177
Db 1816 AACGACCTACCCGAACCTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCC 1757
QY 5178 cgaaggggagaaaaggcggacaggtatccggtaagcggcagggctcggaacacagagagcgcac 5237
Db 1756 CGAAGGGAGAAAGCGGCACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGAGAGCGCAC 1697
QY 5238 gaggaggttccagggggaaacgcctggtatcttttagtctcgtcgtcggggttccgccaact 5297
Db 1696 GAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCTCTGCGGGTTTCGCCACCT 1637
QY 5298 ctgacttgagcgtcgaattttgtgatgctcgtcagggggcggagcctatggaaaaacgc 5357
Db 1636 CTGACTTGAGCGTCGATTTTGTGATGCTGCTCAGGGGGCGGAGCTATGGAAAAACGC 1577
QY 5358 cagcaacgcggccctttttacgggttctcgtgaccttttgcggccttttgcctcacatgttctt 5417
Db 1576 CAGCAACGCGGCCCTTTTACGGTTCTCTGGCCTTTTGTGGCCTTTTGTCTCACATGTTCTT 1517
QY 5418 tcctgcgttatccccctgattctctgtggataacccgtattaccgccttttgcgtgagctgatatc 5477
Db 1516 TCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCCTTTGAGTGAGCTGATAC 1457
QY 5478 cgctgcgcgcagccgaacacgacccgagcgcagcagtcagtcagtcagggaggaagcgggaagcg 5537
Db 1456 CGCTCGCCGACGCCGAACCGACCGAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCG 1397
QY 5538 cccaatacgcacaaacccgcctctccccgcgcgttgcccgattcaatgaagcagctggcacga 5597


```
Db 1396 CCAATACGAAACCGCCTCTCCCGCGCGTGGCCGATTCAATTAATGACAGCTGGCACGA 1337
QY 5598 caggtttcccagctggaagcggcgagcgcaacgcaattaatgagttagctcac 5657
Db 1336 CAGGTTTCCGAGCTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAC 1277
QY 5658 tcattaggcaccgccaggtttacacctttatgcttccggctcgtatgtgtggaattgt 5717
Db 1276 TCATTAGGCACCCAGGCTTTACACTTTATGCTTCGGGCTCGTATGTTGTGGAATTGT 1217
QY 5718 gagcgataacaatttcacacaggaaacagctatgaccatgattacgcca 5767
Db 1216 GAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGGCCA 1167

RESULT 13
U02454
LOCUS U02454 5452 bp DNA circular SYN 29-MAR-1996
DEFINITION Cloning vector pCMVEBNA, complete sequence.
ACCESSION U02454
VERSION U02454.1 GI:413820
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE Cloning vector pCMVEBNA.
JOURNAL Cloning vector pCMVEBNA
REFERENCE artificial sequence; vectors.
AUTHORS 1 (bases 1 to 5452)
Kitts,P.A.
CLONTECH Vectors On Disc version 1.3
2 (bases 1 to 5452)
Unpublished
Swirski,R.A., Van Den Berg,D., Murphy,A.J., Lambert,C.M.,
Friedberg,E.C. and Schimke,R.T.
TITLE Improvements in the Epstein-Barr-based shuttle vector system for
JOURNAL direct cloning in human tissue culture cells
REFERENCE Methods: A Companion to Methods in Enzymology 4, 133-142 (1992)
AUTHORS 3 (bases 1 to 5452)
Kitts,P.A.
Direct Submission
Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
This vector can be obtained from CLONTECH Laboratories, Inc., 1020
East Meadow Circle, Palo Alto, CA 94303, USA. To place an order
call (415) 424-8222 or (800) 662-2566, extension 1. International
customers, please contact your local distributor. For technical
information, call (415) 424- 8222 or (800) 662-2566, extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH; this vector has not been
completely sequenced. If you suspect there is an error in this
sequence, please contact CLONTECH's Technical Service Department at
(415) 424-8222 or (800) 662-2566, extension 3 or E-mail
TECH@CLONTECH.COM.
FEATURES
source Location/Qualifiers
1. .5452
/organism="Cloning vector pCMVEBNA"
/db_xref="taxon:31797"
BASE COUNT 1362 a 1246 c 1736 g 1108 t
ORIGIN
Query Match 45.6%; Score 2630; DB 12; Length 5452;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3137 cactggccgtcgttttacaaacgtcgtgactgggaaaccctggcgttaccaccaacttaac 3196
Db 2823 CACTGGCCGTCGTTTACAACGTCGTGACTGGGAAACCCTGGCGTTACCCAACTTAATC 2882
QY 3197 gccttgacacatcccccttttcgccagctggcgtaataagcgaagagccgcaccgac 3256
Db 2883 GCCTTGACGACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC 2942
QY 3257 gcccttcccaacagttgcgcagcctgaatggcgaatggcgctgatgggtatttctcc 3316
```

```
Db 2943 GCCCTTCCCAACAGTTGGCGAGCCTGAATGGCAATGGCGCCTGATGGGCTATTTCTCC 3002
QY 3317 ttacgcatctgtcggttatattcacaccgcataatggtgcactctcagtaacaatctgctctg 3376
Db 3003 TTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGCTCTG 3062
QY 3377 atgccgcatagttaagccagccccgcacacccgccaaacacccgcctgacgcgcctgacggg 3436
Db 3063 ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCCGCTGACGCCCTGACGGG 3122
QY 3437 cttgtctgctcccgccatccgcttacagacaaagctgtgacctctccgggagctgcatgt 3496
Db 3123 CTTGTCTGCTCCCGCATCCGCTTACAGACAAAGCTGTGACCGTCTCCGGGAGCTGCATGT 3182
QY 3497 gtcagagggtttcacccgctcatcaccgaaacgcgcgagacgaaagggcctcgtgatacggc 3556
Db 3183 GTCAGAGGTTTTCACCGTCTATCACCGAAACCGCGGAGACGAAAGGCCCTCGTATACGCC 3242
QY 3557 tatttttataggttaatgtcatgataataatggttttcttaaatacattcaaatatgtatc 3616
Db 3243 TATTTTATAGGTTAATGTATGATAATAATAGTTCCTTAGACGTGAGTGGCACTTTTC 3302
QY 3617 ggggaaatgtgcgggaacccctattgtttatttttctaaatacattcaaatatgtatc 3676
Db 3303 GGGGAAATGTGCGGGAACCCCTATTGTGTTTATTTTCTAAATACATTCAAAATATGTATC 3362
QY 3677 cgctcatgagacaataaacccctgataaaatgcitcaataataattgaaaaaaggaagatga 3736
Db 3363 CGCTCATGAGACAATAACCTTGATAAATGCTTCAATAATATTGAAAAAGGAGAGATGA 3422
QY 3737 gtattcaacatttcctcgtgtcgccctattcccttttttgcggcatttttgccttccctgttt 3796
Db 3423 GTATTCAACATTTCCGTGTGCGCCCTTATTTCCCTTTTTCGGGCATTTTGGCCTTCCTGTTT 3482
QY 3797 ttgctcaccacgaacgcgtggtgaaagtaaaagatgctgaagatcagttgggtgcacgag 3856
Db 3483 TTGCTCACCCAGAAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 3542
QY 3857 tgggttacatcgaactggatctcaacagcggtaagatccttgagagtttgcgccccgaag 3916
Db 3543 TGGGTTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAG 3602
QY 3917 aacgttttccaatgatgagcacttttaaagtcttctgctatgtggcgcgttattatcccgta 3976
Db 3603 AACGTTTTCCAATGATGAGCACTTTTAAAGTCTGCTATGTGGCGCGGTATATCCCGTA 3662
QY 3977 ttgacgcccgggcaagagcaactcggctgcgcgcatacacactattctcagaatgacttggttg 4036
Db 3663 TTGACGCCCGGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAAATGACTTGGTTG 3722
QY 4037 agtactcaccagtcacagaaaaagcatcttaacggatggcatgacagtaagagaattatgca 4096
Db 3723 AGTACTCACCGCTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATGCA 3782
QY 4097 gtgctgccataaccatgagtataaacactgcggccaaacttacttctgacaacgacgcgag 4156
Db 3783 GTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAG 3842
QY 4157 gaccgaaggagtaaccgcttttttgcaacaatgggggatcatgtaactgcgcttgatc 4216
Db 3843 GACCGAAGGAGCTAAACCGCTTTTTCACAAACATGGGGATCATGTAACTCGCCTTGATC 3902
QY 4217 gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacacgacgatgcctg 4276
Db 3903 GTTGGGAACCCGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACACGATGCCTG 3962
QY 4277 tagcaatggcaacaacgcttgcgcaaaactattaaactggcgaactacttactctagcttccc 4336
Db 3963 TAGCAATGGCAACAACGTTGGGCAAACTATTAACTGGCGAACTACTTACTTACTGCTTCCC 4022
QY 4337 ggcaacaattaatagactggatggagcggaataaagttgcaggaccacttctgcgctcgg 4396
```


Db 6541 ACTGGCCGTCGCTTTTACAACGTCGTGACTGGGAAAAACCCCTGGCGTTACCCAACTTAATCG 6482

Qy 3198 ccttgacgacatcccccttctgcagctggcgtaatagcgaagagggcccgacccgatcg 3257

Db 6481 CCTTGCAGACATCCCCCTTTCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCG 6422

Qy 3258 cccttcccaacagttgcgcagccctgaatggcgccctgatgcgggtatttctcct 3317

Db 6421 CCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCCCTGATGCGGTATTTTCTCCT 6362

Qy 3318 tacgcatctgtgcggtatttcacacccgcataatggtgcactctcagtacaaatctgctctga 3377

Db 6361 TACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGA 6302

Qy 3378 tgcgcatagtttaagccagccccgcacaccccgccaaacaccccgctgaacgcgcctgacgggc 3437

Db 6301 TGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGGCCCCCTGACGGGC 6242

Qy 3438 ttgtctgctcccgcatccgcttacagacaagctgtgacccgtctccgggagctgcattg 3497

Db 6241 TTGCTGTCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCTGCATGTG 6182

Qy 3498 tcagaggttttcacccgtcatcacgaacgcgcgagacgaagggcctcgtgatacgcct 3557

Db 6181 TCAGAGGTTTTCACCGTCATCACCGAACCGCGGAGACGAAAGGGCCTCGTGATACGCCT 6122

Qy 3558 atttttataggttaatgtcatgataataatggttctcttagacgtcagtggtgcacttttcg 3617

Db 6121 ATTTTATAGGTTAATGTCAATGATAATAATGTTTCTTAGACGTGAGGTGGCACTTTTCG 6062

Qy 3618 gggaaatgtgcgcggaaacccctatttgttattttctctaaatacatcctcaaatatgtatcc 3677

Db 6061 GGGAAATGTGCGCGGAACCCCTATTGTTTATTATTTTCTAAATACATTCAAATATGTATCC 6002

Qy 3678 gctcatgagacaaataaccctgataaataatgcttcaataattgaaaaaggaagatgag 3737

Db 6001 GCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATATGAG 5942

Qy 3738 tatteacaattccgtgtcgcccttattcccttttttgcggcatttgccttcctgtttt 3797

Db 5941 TATTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTCGCGCATTTTGCCCTTCCTGTTTT 5882

Qy 3798 tgctcaccagaaaacgctggtgaaagtataaagatgctgaagatcagttgggtgcacgagt 3857

Db 5881 TGCTCACCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT 5822

Qy 3858 gggttacatcgaactggatctcaacagcggtaagatccttgagagtttctgccccgaaga 3917

Db 5821 GGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGA 5762

Qy 3918 acgtttttccaatgatgagcactttttaaagtctgctatgtggcgcggtattatccccgtat 3977

Db 5761 ACGTTTTCGAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTAT 5702

Qy 3978 tgacgcggggcaagagcaactcggtcgccgcatacacataattctcagaatatgacttggtga 4037

Db 5701 TGACGGCGGCAAGAGCAACTCGGTCCCGGCATACACTATTCTCAGAAATGACTTGGTTGA 5642

Qy 4038 gtactcaccagttcacagaaaaagcatcttacggatggcatgacagtacagtaagagaattatgcag 4097

Db 5641 GTACTACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAG 5582

Qy 4098 tgctgccataaccatgagtataacactgcggcccaacttacttctcgacaacgatcggagg 4157

Db 5581 TGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGGAGG 5522

Qy 4158 accgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactgccttgatcg 4217

Db 5521 ACCGAAGGAGCTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCCTTGATCG 5462

Qy 4218 ttgggaaacggagctgaatgaagccataccaaaacgacgagcgtgcacaccacgatgcctgt 4277

Db 5461 TTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGGTGACACCCACGATGCCTGT 5402

Qy 4278 agcaatggcaacaacgcttgcgaaaaactattaaactggcgaactacttactctagcttcccc 4337

Db 5401 AGCAATGGCAACAACAGTTGCGCAAACTATTAACTGGCGAACTACTTACTTAGCTTCCCG 5342

Qy 4338 gcaacaaattaatagactggatggagggcgataaaagtgcaggaccacttctgcgctcgcc 4397

Db 5341 GCAACAATTAATAGACTGGATGGAGGCGGATAAAAGTTGCAGGACCACCTTCTCGCCTCGG 5282

Qy 4398 ccttcccgctggctggtttattgctgataaaatctggagccggtgagcgggtctcgcgg 4457

Db 5281 CCTTCCGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTAGCGTGGTCTCGCGG 5222

Qy 4458 tatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgac 4517

Db 5221 TATCATTCAGCACTGGGGCCAGATGGTAAGCCCTCCCCTATCGTAGTTATCTACACGAC 5162

Qy 4518 ggggagtcaggcaactatggatgaacgaataagacagatcgctgagataggtgcctcact 4577

Db 5161 GGGGAGTCAGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT 5102

Qy 4578 gattaagcattggttaactgtcagaccaagtcttactcataatacttagattgatttaa 4637

Db 5101 GATTAAGCATTGGTAACCTGTAGACCAAGTTTACTCATATATACTTTAGATTGATTAAA 5042

Qy 4638 acttcatttttaatttaaaggatctaggtaagatcccttttgataatctcatgaccaa 4697

Db 5041 ACTTCATTTTAAATTTAAAAGGATCTAGTGAAAGATCCCTTTTGATAATCTCATGACCA 4982

Qy 4698 aatcccttaacgtgagtttctcgtccactgagcgtcagaccccgtagaaaaagatcaagg 4757

Db 4981 AATCCCTTAACGTGAGTTTTCGTTCCTACTGAGCGCTCAGACCCCGTAGAAAAAGATCAAAGG 4922

Qy 4758 atcttcttgagatcccttttctgcgctaactctgctgtgcatacaaaaaaacacc 4817

Db 4921 ATCTTCTTGAGATCCTTTTTCGTGCGCGTAATCTGCTGTGCTTGCACAAACAAAAACCCACC 4862

Qy 4818 gctaccagcgggtggtttgttgcgggatacaagagctaccaactcttttccgaaggtaac 4877

Db 4861 GCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAAC 4802

Qy 4878 tggcttcagcagagcgcagatataccaaataactgtccttctagtgtagcgtagttaggcca 4937

Db 4801 TGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTTAGTGTAGCCGTAGTTAGGCCA 4742

Qy 4938 ccacttcaagaactctgtagcaacgcctacatacctcgtctgctaactcctgttaccagt 4997

Db 4741 CCACITCAAGAACTCTGTAGCACCGCCTACATAACCTCGCTCTGCTTAATCTGTACCAGT 4682

Qy 4998 ggctgctgcagtggcgataaagtctgtcttaccgggttggaactcaagacgatatgtacc 5057

Db 4681 GGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGTTGGACTCAAGACGATAGTTACC 4622

Qy 5058 ggataaaggcgacgagcggtcgggctgaacgggggttcgtgcacacagcccagcttgagcgc 5117

Db 4621 GGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCTGTGCACACAGCCCGCTGGAGCG 4562

Qy 5118 aacgacctacaccgaactgagataacctacagcgtgagctatgagaaaagcgccacgcttcc 5177

Db 4561 AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCCCGACGCTTCC 4502

Qy 5178 cgaagggagaaagcggaaggtatccggtaagcggcaggggtcggaacagagagcgcac 5237

Db 4501 CGAAGGGAGAAAGCGGACAGGTATCCCGTAAGCGGACGGGTGCGGAACAGGAGAGCGCAC 4442

Qy 5238 gagggagcttccagggggaaacgcctggtatcttatagtcctgtcgggtttcgcacct 5297

Db 4441 GAGGGAGCTTCCAGGGGGAACCGCCTGSTATCTTATAGTCTCTGCTGGGTTTCGCCACCT 4382

Qy 5298 ctgacttgagcgtcgatttttctgtagtctcgtcaggggggcgagcctatggaaaaaacgc 5357

Db 4381 CTGACTTGACGCTCGATTTTTGTGTGATGCTCTGTCAGGGGGGGGAGCCCTATGGAAAAACGC 4322


```
QY 5358 cagcaacgcgcctttttacgggttcctcggttcctgttgcccttttgcggccttttgcctcacatgttctt 5417
      |||
Db 4321 CAGCAACGCGGCTTTTACGGTTCTCTGGCTTTTGTGGCCTTTTGTCTCACATGTTCTT 4262
      |||
QY 5418 tcctgcgttatccctgatctctgtggataaccggtattaccgccttttgagtgagctgatac 5477
      |||
Db 4261 TCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC 4202
      |||
QY 5478 cgctgcgcgagccgaacgacgagcgagcgagtcagtgagcgaggaagcgagcg 5537
      |||
Db 4201 CGCTCGCCGAGCCGAACGACCGAGCGAGCGAGTCAGTGAGCGAGGAAGCGGAGCG 4142
      |||
QY 5538 cccaatacgaacacgcctctctcccgcggttgccgattcattaatgagctggcacga 5597
      |||
Db 4141 CCCNATACGNAACCGCCTCTCCCCGCGCTGGCCGATTCAATTAATGAGCTGGCACGA 4082
      |||
QY 5598 caggtttccgactggaagcgggcagtgagcgcaacgcaatgaatgagtgagctcac 5657
      |||
Db 4081 CAGGTTTCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCAC 4022
      |||
QY 5658 tcattaggcaccacaggcctttacacatttatgcttcggcgtcgatgttggtggaattgt 5717
      |||
Db 4021 TCATTAGGCACCCAGCGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGT 3962
      |||
QY 5718 gagcggataacaatttcacacaggaacacgctatgaccatgattacgcca 5767
      |||
Db 3961 GAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGATTACGCCA 3912
      |||

RESULT 15
XXU02435/C
LOCUS XXU02435 6894 bp DNA circular SYN 29-JAN-1997
DEFINITION Cloning vector pSVbeta, complete sequence.
ACCESSION U02435
VERSION U02435.1 GI:413801
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS MacGregor,G.R. and Caskey,C.T.
TITLE Construction of plasmids that express E. coli beta-galactosidase in mammalian cells
JOURNAL Nucleic Acids Res. 17 (6), 2365 (1989)
MEDLINE 89202057
REFERENCE 2 (bases 1 to 6894)
AUTHORS Kitts,P.A.
TITLE CLONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 6894)
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This vector can be obtained from CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA. To place an order call (415) 424-8222 or (800) 662-2566, extension 1. International customers, please contact your local distributor. For technical information, call (415) 424- 8222 or (800) 662-2566, extension 3. This sequence has been compiled from information in the sequence databases, published literature and other sources, together with partial sequences obtained by CLONTECH; this vector has not been completely sequenced. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.
FEATURES
source Location/Qualifiers
1. .6894
/organism="Cloning vector pSVbeta"
/db_xref="taxon:31847"
BASE COUNT 1630 a 1780 c 1830 g 1654 t
ORIGIN
```

```
Query Match 45.6%; Score 2630; DB 12; Length 6894;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3138 actggccgctggtttttacaacgtcgtgactgggaaaaacccctggcgtttaccacaacttaatcg 3197
      |||
Db 6894 ACTGGCCGTCGTTTACAACTGCTGACTGGGAAAAACCTGGCGTTACCCAACCTAATCG 6835
      |||
QY 3198 ccttgccagcacatcccccttccgcccagctggcgtaataagcgaagagggcccgaccgatcg 3257
      |||
Db 6834 CCTTGCAGCACATCCCCCTTTCCGCCAGCTGCGCTAATAGCGAAGAGGCCCGCACCGATCG 6775
      |||
QY 3258 cccttcccaacagttgcgcagcctgaaatggcgaaatggcgctgatgcggtattttctcct 3317
      |||
Db 6774 CCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTGATGCGGTATTTTCTCCT 6715
      |||
QY 3318 tacgcatctgtgcggtatttcacacccgcataatggtgcactctcagtaacaatctgctctga 3377
      |||
Db 6714 TACGCATCTGTGGGTATTTTACACCCGCATATGTTGTCACCTCTCAGTACAATCTGCTCTGA 6655
      |||
QY 3378 tgccgcatagttaagccagccccgcacacccgcaacacccgcctgacggccctgacgggc 3437
      |||
Db 6654 TGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGGCCCTGACGGGC 6595
      |||
QY 3438 ttgtctgctcccggcacccgcttacagacaagctgtgacccgtctccgggagctgcatgtg 3497
      |||
Db 6594 TTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCCGTCTCCGGGAGCTGCATGTG 6535
      |||
QY 3498 tcagagggttttcaccgtcatcacccgaaaaacgcgcgcgcgcgcgcgcgcgcgcgcgcct 3557
      |||
Db 6534 TCAGAGGTTTTCACCGTCAATCACCGAAAAACGCGCGAGACGAAAGGCCCTCGTGATACGCCT 6475
      |||
QY 3558 atttttataggttaatgtcatgataataatggttttcttagacgtcagtggcacttttcg 3617
      |||
Db 6474 ATTTTATAGGTTAATGTCAATGATAATAATAGGTTTCTTAGACGTGAGTGGCACTTTTCG 6415
      |||
QY 3618 gggaaatgtgcgggaaacccctatttgtttatttttctaaatacattcaaatatgtatcc 3677
      |||
Db 6414 GGGAAATGTGCGGGAAACCCCTATTGTTTATTTTCTAAATACATTCAAAATATGTATCC 6355
      |||
QY 3678 gctcatgagacaataaaccttgataaaatgcttcaataataattgaaaaaggaagatgatgag 3737
      |||
Db 6354 GCTCATGAGACAATAACCTTGATAAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAG 6295
      |||
QY 3738 tattcaacatttccggtgctgcccttattcccttttttgcggcattttgccttcctgtttt 3797
      |||
Db 6294 TATTCAACATTTCCGTGTGCCCTTATTCCCTTTTTCGGGCATTTTTCCTTCTCTGTTT 6235
      |||
QY 3798 tgctcaccacgaacgcgtggtgaaagtaaaagatgctgaagatcagttgggtgcacgagt 3857
      |||
Db 6234 TGCTCACCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT 6175
      |||
QY 3858 gggttacatcgaactggatctcaacacgcggttaagatccttgagagtttccgcccgaaga 3917
      |||
Db 6174 GGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGA 6115
      |||
QY 3918 acgttttccaatgatgagcacttttaagtctctgtatgtggcgcgttattatccccgtat 3977
      |||
Db 6114 ACGTTTTCNAATGATGAGCACTTTTAAAGTCTGCTATGTGGCGCGGTATATCCCGTAT 6055
      |||
QY 3978 tgacgcgggcaagagcaactcggtcgcgcgcacatacacactattctcagaatgacttggtga 4037
      |||
Db 6054 TGACGCCGGGCAAGAGCAACTCGGTCCGCCCATACACTATTCTCAGATGACTTGGTTGA 5995
      |||
QY 4038 gtactcaccagtcacagaaaaagcatcttaaggatggcatgacagtaagagaattatgcag 4097
      |||
Db 5994 GTACTCACCAGTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATGCAG 5935
      |||
QY 4098 tgctgccataaccatgagtgataaacactcgggccaacttacttcttgacaacgatcggagg 4157
      |||
Db 5934 TGCTGCCATAACCATGAGTGATAAACACTGCGGCCCAACTTACTTCTTGACAACGATCGGAGG 5875
      |||
```


QY	4158	accgaaggagtaaccgccttttttcacaacatgggggatcatgtaactcgccttgatcg	4217
Db	5874	ACCGAAGGAGCTAACCGCTTTTGTGCAACATGGGGATCATGTAACTGCGCTTGATCG	5815
QY	4218	ttgggaaccggagctgaatgaagccataccaaaacgacgagcgtgacaccacgatgcctgt	4277
Db	5814	TTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACGATGCCTGT	5755
QY	4278	agcaatggcaacaacggttgcgcaaaactattaaactggcgaactacttactctagcttcccg	4337
Db	5754	AGCAATGSCAACAAACGTTGCGCAACATATTAACTTGCCGAACACTACTACTCTAGCTTCCCG	5695
QY	4338	gcaacaattaatagactggatggaggggataaaagtgcaggagcaacttctgcgctggc	4397
Db	5694	GCAACAATTAATAGACTGGATGGAGGGCGGATAAAGTTGCAGGACCACCTTCTGCGCTCGGC	5635
QY	4398	ccttccggctggctggtttatttgctgataaaatctggagccggtgagcgggtggtctcgcgg	4457
Db	5634	CTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGTTGAGCCGTTGCGGTCTCGCGG	5575
QY	4458	tatcattgcagcactggggccagatggatgaacgaatagacagatcgctgagatggtgcctcact	4517
Db	5574	TATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC	5515
QY	4518	ggggagtcaggcaactatggatgaacgaatagacagatcgctgagatggtgcctcact	4577
Db	5514	GGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGTGCCTCACT	5455
QY	4578	gattaagcaattggtaaactgtcagaccaaagtttactcatatatatacttttagattgatttaa	4637
Db	5454	GATTAAGCATTTGGTAACCTGTCAGACCCAAGTTTACTCATATATACTTTAGATTGATTAAA	5395
QY	4638	acttcatttttaatttaaaaggatctagggtgaagatcccttttttgataaatctcatgaccaa	4697
Db	5394	ACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCATGACCAA	5335
QY	4698	aatcccttaacgtgagttttctgttcactgagcgtcagacccccctagaaaaagatcaaaag	4757
Db	5334	AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAAG	5275
QY	4758	atcttcttgagatccctttttctgcggtaatctgctgcttgcaaaacaaaaaacacc	4817
Db	5274	ATCTTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACCAACC	5215
QY	4818	gctaccagcgggtggttgttgccggatcaagagctaccaaactctttttccgaaggtaac	4877
Db	5214	GCTACCAAGCGGTGGTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC	5155
QY	4878	tggcttcacgacgagcgcagataccacaatactgtccttctagtagtagccgtagttaggcca	4937
Db	5154	TGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCCGTAGTTAGGCCA	5095
QY	4938	ccacttcagaactctgtagcaccgcctacataacctgcctctgctaactctgttaccagt	4997
Db	5094	CCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATTCCTGTTACCCAGT	5035
QY	4998	ggctgctgcagtgggcgataagtcgttcttaccgggttggaactcaagacgatatgttacc	5057
Db	5034	GGCTGCTGCAGTGGCGGATAAAGTCGTGCTTACCGGTTGGACTCAAGACGATAGTTACC	4975
QY	5058	ggataaggcgcagcggctcggggtgaacgggggggttcgtgcacacagcccgcttggagcg	5117
Db	4974	GGATAAGGCGCAGCGGTGCGGCTGAACGGGGGTTTCGTGCACACACGCCAGCTTGGAGCG	4915
QY	5118	aacgacctacaccgaactgagatcacctacagcgtgagctatgagaaaaagccacgcttcc	5177
Db	4914	AACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCGCCACGCTTCC	4855
QY	5178	cgaagggagaaaagcggacagggtatccggttaagcggcagggctcggaacaggagagcgcac	5237
Db	4854	CGAAGGGAGAAAGCGGCACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCAC	4795
QY	5238	gagggagcttccaggggggaaacgcctgggtatctttatagtcctgctcggggttctgccacct	5297

Db	4794	GAGGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCCTGCGGGTTTCGCCACCT	4735
QY	5298	ctgacttgagcgtcgcgatttttgtgatctcgcgcagggggcggagcctatggaaaaaacgc	5357
Db	4734	CTGACTTGAGCGTTCGATTTTGTGATGCTCGTCAAGGGGCGGAGCCTATGGA AAAACGC	4675
QY	5358	cagcaaacggcgccttttttacggttcctggccttttgcggccttttgctcacatgttctt	5417
Db	4674	CAGCAACGGGCGCTTTTACGGTTCCCTGGCTTTTGTGGCCTTTTGCTCACATGTTCTT	4615
QY	5418	tcctgcgttatccccctgattctgtgtgataaacctgattaccgcctttgagtgagctgatac	5477
Db	4614	TCCTGCGTTATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTTGAGTGAGCTGATAC	4555
QY	5478	cgctcgcgcagccgaaacgacccgagcgcagcagtcagtgagcgaagaaagcgaagcgcg	5537
Db	4554	CGCTCGCCCGCAGCCGAACGACCCGAGCCGCGAGTCAGTGACCGGAGGAAGAGCGG	4495
QY	5538	cccaatacacaaacccgcctctctccccgcggttgcccgattcattaatgcagctggcacga	5597
Db	4494	CCCAATACGCAAAACCGCCTCTCCCCCGCGTTGGCCGATTTCATTAATGTCAGCTGGCACGA	4435
QY	5598	caggtttcccgactggaaaagcgggcagtgagcgaacgcgaatattagttagctcac	5657
Db	4434	CAGGTTTCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAC	4375
QY	5658	tcattaggcaacccagggtttacacttttatgcttccggctcgtagttgtgtggaattgt	5717
Db	4374	TCATTAGGCACCCAGGCTTTACACTTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGT	4315
QY	5718	gagcggataacaatttcacacacaggaacacagctatgaccatgattacgcca	5767
Db	4314	GAGCGGATAACAATTTCACACAGGAACACAGCTATGACCATGATTACGCCA	4265

Search completed: August 31, 2002, 22:41:47
Job time: 25536 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:58:22 ; Search time 824.51 Seconds
(without alignments)
3592.045 Million cell updates/sec

Title: US-09-810-861b-5
Perfect score: 1725
Sequence: 1 atgaggccccgcagtgctt.....aattgtcagcgccacctga 1725

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1725	100.0	1725	24	AA17549
2	1722	99.8	1845	24	AA17493
3	1722	99.8	2256	16	AAQ99002
4	1712	99.2	5767	24	AA17547
5	1712	99.2	14446	24	AA17548
6	1252	72.6	2253	11	AAQ05998
7	1070	62.0	12113	24	AA17492
8	489	28.3	1800	11	AAQ05999
9	35	2.0	35	17	AAT34118
					Synthetic human ac
					Human cDNA encoding
					Human acetylcholin
					Plasmid pTM034. S
					Plasmid pTM036. S
					Sequence encoding
					Human acetylcholin
					Sequence encoding
					Acetylcholinestera

10	33	1.9	2089	24	AB199790	Mouse ischaemic co
11	28	1.6	8418	22	AA541906	Genomic sequence #
12	27	1.6	816	23	AA574821	DNA encoding novel
13	27	1.6	1532	23	AA574822	DNA encoding novel
14	27	1.6	1538	22	AA09548	Human protease pro
15	25	1.4	496	23	AA574456	DNA encoding novel
16	25	1.4	2523	22	AA574457	Human MBSP4 polype
17	25	1.4	4975	23	AA574457	DNA encoding novel
C 18	22	1.3	4173	21	AA62899	Murine JNK3 blindin
C 19	22	1.3	4200	21	AA62900	Murine JNK3 blindin
C 20	22	1.3	4266	21	AA62902	Murine JNK3 blindin
C 21	22	1.3	4269	21	AA62901	Murine JNK3 blindin
C 22	21	1.2	21	24	AA517550	Human acetylcholin
C 23	20	1.2	20	19	AA54213	Nuclease resistant
C 24	20	1.2	20	19	AA54214	Nuclease resistant
C 25	20	1.2	20	19	AA54215	Nuclease resistant
C 26	20	1.2	20	19	AA541281	Human AChE mRNA sp
C 27	20	1.2	20	19	AA541282	Human AChE mRNA sp
C 28	20	1.2	20	19	AA541283	Human AChE mRNA sp
C 29	20	1.2	20	22	AA44814	Antisense oligonuc
C 30	20	1.2	20	22	AA44815	Antisense oligonuc
C 31	20	1.2	20	22	AA44817	Antisense oligonuc
C 32	20	1.2	1698	20	AA86613	CDNA encoding an a
C 33	20	1.2	2369	23	ABL13421	Drosophila melanog
C 34	20	1.2	2478	20	AA86614	CDNA encoding an a
C 35	20	1.2	5124	23	AA595024	DNA encoding human
C 36	20	1.2	6900	23	AA595022	DNA encoding murin
C 37	20	1.2	6901	23	AA595027	DNA encoding murin
C 38	20	1.2	7020	24	AA26312	Human G-protein co
C 39	20	1.2	7119	23	AA595023	DNA encoding murin
C 40	20	1.2	7119	23	AA595025	DNA encoding murin
C 41	20	1.2	7172	23	AA595026	DNA encoding human
C 42	20	1.2	7694	22	ABA07755	Human ovarian and
C 43	20	1.2	7694	22	AA03544	Human reproductive
C 44	20	1.2	18413	23	ABL13420	Drosophila melanog
C 45	19	1.1	19	24	AA517551	Human acetylcholin

ALIGNMENTS

RESULT 1	
AA517549	
ID	AA517549 standard; DNA; 1725 BP.
XX	
AC	AA517549;
XX	
DT	25-FEB-2002 (first entry)
XX	
DE	Synthetic human acetylcholinesterase gene.
XX	
KW	Human; acetylcholinesterase; AChE; antidote; pesticide;
KW	transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW	muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW	organophosphate(OP)-modified AChE; pyridostigmine bromide; ds.
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	WO200171014-A2.
XX	
PD	27-SEP-2001.
XX	
PF	16-MAR-2001; 2001WO-US08468.
XX	
PR	17-MAR-2000; 2000US-190440P.
XX	
PA	(MORT/) MOR T.
PA	(SORE/) SOREQ H.
PA	(ARNT/) ARNTZEN C.
PA	(MASO/) MASON H.
PA	(BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX	

Db 1561 aatgagccccgagaccccaaggccccacaaatggccccgtacacggcggggctcagcag 1620
QY 1621 tacgttagtctggaacctgcggcccgctggaggtgcggcggggctgcgcgccaggcctgc 1680
Db 1621 tacgttagtctggaacctgcggcccgctggaggtgcggcggggctgcgcgccaggcctgc 1680
QY 1681 gccttctggaaccgcttctctcccaaaattgctcagcgccacctga 1725
Db 1681 gccttctggaaccgcttctctcccaaaattgctcagcgccacctga 1725
RESULT 2
AAS17493
ID AAS17493 standard; cDNA; 1845 BP.
XX
AC AAS17493;
XX
DT 26-FEB-2002 (first entry)
XX Human cDNA encoding acetylcholinesterase, ACHE.
DE
XX Human; ss; ACHE; acetylcholinesterase; polymorphic variant; haplotyping;
KW genotyping; neurological disease; Parkinson's disease;
KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1845
FT /*tag= a
FT /product= "ACHE"
FT replace (36,T)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT replace (101,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT replace (183,C)
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT replace (838,G)
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT replace (873,T)
FT /*tag= f
FT /standard_name= "Single nucleotide polymorphism"
FT replace (1032,A)
FT /*tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT replace (1057,A)
FT /*tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT replace (1431,T)
FT /*tag= i
FT /standard_name= "Single nucleotide polymorphism"
XX
PN WO200179219-A2.
XX
PD 25-OCT-2001.
XX
PF 11-APR-2001; 2001WO-US11853.
XX
PR 14-APR-2000; 2000US-197173P.
XX (GENA-) GENAISSANCE PHARM INC.
PA (KAZE/) KAZEMI A.
XX Bentivegna SC, Chew A, Choi JY, Koshy B;
XX WPI; 2002-055248/07.
DR P-PSDB; AAU11231.
XX

PT New polymorphic variants comprising acetylcholinesterase (ACHE) isogene, useful in expressing ACHE protein for use in screening for candidate drugs to treat diseases related to ACHE activity, e.g. neurological diseases or cancer -
XX
PS Claim 26; Fig 2; 79pp; English.
XX
CC The invention relates to a polynucleotide comprising a polymorphic variant of an acetylcholinesterase (ACHE) gene or fragment, protein or complement, the variant comprising an ACHE isogene defined by a haplotype selected from haplotypes 1-20 listed in the specification. Also included are methods for haplotyping and genotyping the ACHE gene of an individual, a method for predicting a haplotype pair for the ACHE gene of an individual, a method for identifying an association between a trait and at least one haplotype or haplotype pair of ACHE gene, recombinant nonhuman organisms transformed or transfected with the polynucleotide where the organism expresses ACHE protein encoded by the first nucleotide sequence or encoded by the polymorphic variant sequence, an isolated antibody specific for and immunoreactive with ACHE, a method of screening for drugs targeting the polypeptide contacting ACHE polymorphic variant with a candidate agent and assaying for binding activity, a computer system for storing and analysing polymorphism data for ACHE gene and a genome anthology for ACHE gene which comprises ACHE isogenes defined by haplotypes 1-20 given in the specification.
CC The polymorphisms are useful for studying the biological function of ACHE as well as in identifying drugs targeting this protein for the treatment of disorder related to its abnormal expression or function.
CC The polymorphic variants may also be used in screening for compounds targeting ACHE to treat a specific condition or disease predicted to be associated with ACHE activity e.g. neurological diseases (e.g. Parkinson's disease and Alzheimer's disease), cancer, leukaemia, and tumours. The ACHE gene maps to human chromosome 7q22. The present sequence is the coding sequence of the ACHE gene.

SQ Sequence 1845 BP; 302 A; 590 C; 608 G; 345 T; 0 other;

Query Match 99.8%; Score 1722; DB 24; Length 1845;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagggccccgcagtgctgtctgtgcacacgccttccctggcttccccactctctcctc 60
Db 1 atgagggccccgcagtgctgtctgtgcacacgccttccctggcttccccactctctcctc 60
QY 61 ctctctgtgctcctggtgggagagtggtgggctgagggcgaggagatgcagagctgctg 120
Db 61 ctctctgtgctcctggtgggagagtggtgggctgagggcgaggagatgcagagctgctg 120
QY 121 gtgacggtgcgtggggccgctgcggggcattgcctgaagacccccggggccctgctc 180
Db 121 gtgacggtgcgtggggccgctgcggggcattgcctgaagacccccggggccctgctc 180
QY 181 tetgctttcctgggcatcccccttgcggagccacccatgggacccccgcgttcttgc 240
Db 181 tetgctttcctgggcatcccccttgcggagccacccatgggacccccgcgttcttgc 240
QY 241 ccggagcccaagcagccttggtcaggggtggtagacgtacacaccttccagagtgtctgc 300
Db 241 ccggagcccaagcagccttggtcaggggtggtagacgtacacaccttccagagtgtctgc 300
QY 301 taccaatatgtggacacccctatccccagggttttggggcaccgagatgtggaaccccaac 360
Db 301 taccaatatgtggacacccctatccccagggttttggggcaccgagatgtggaaccccaac 360
QY 361 cgtgagctgagcagaggactgcctgtacctcaacgtgtggacaccataccccccggcctaca 420
Db 361 cgtgagctgagcagaggactgcctgtacctcaacgtgtggacaccataccccccggcctaca 420
QY 421 tccccacccccctgtcctcgttgatctatgggggtgtggttctctacagtggggcctctcc 480
Db 421 tccccacccccctgtcctcgttgatctatgggggtgtggttctctacagtggggcctctcc 480

QY 481 ttggacgtgtacgatggccgcttctcttggtacagccgagaggactgtgctggtgccatg 540
Db 481 ttggacgtgtacgatggccgcttctcttggtacagccgagaggactgtgctggtgccatg 540
QY 541 aactacccgggtggagacctttggcttctcctggccctgccggggagccgagagcccccggc 600
Db 541 aactacccgggtggagacctttggcttctcctggccctgccggggagccgagagcccccggc 600
QY 601 aatgtgggtcctcctggatcagaggtcgccctgcagtggtgcaggagaacgtggcagcc 660
Db 601 aatgtgggtcctcctggatcagaggtcgccctgcagtggtgcaggagaacgtggcagcc 660
QY 661 ttccgggggtgacccgacatcagtgacgctgtttggggagagcgcggggagccgctcggtg 720
Db 661 ttccgggggtgacccgacatcagtgacgctgtttggggagagcgcggggagccgctcggtg 720
QY 721 ggcattgcacctgctgtccccgccccagcggggcccctgtttccacagggccgtgctgcagagc 780
Db 721 ggcattgcacctgctgtccccgccccagcgggcccctgtttccacagggccgtgctgcagagc 780
QY 781 ggtgcccccaaatggacctggggccagcggtgggcatgggagagggcccgctgcagggccacg 840
Db 781 ggtgcccccaaatggacctggggccagcggtgggcatgggagagggcccgctgcagggccacg 840
QY 841 cagctggccccaccttgtgggctgtcctccagggcgacactggtgggaatgacacagagctg 900
Db 841 cagctggccccaccttgtgggctgtcctccagggcgacactggtgggaatgacacagagctg 900
QY 901 gtacgctgccttcggaacagaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
Db 901 gtacgctgccttcggaacagaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
QY 961 cctcaagaagcgtcttcggttctcctgcctgcctggtgtagatggagacttcctcagt 1020
Db 961 cctcaagaagcgtcttcggttctcctgcctgcctggtgtagatggagacttcctcagt 1020
QY 1021 gacaccccgagggccctcatcaacggggagacttccacggccctgcagtggtgggt 1080
Db 1021 gacaccccgagggccctcatcaacggggagacttccacggccctgcagtggtgggt 1080
QY 1081 gtggtgaaggatgagggctcgatatttctggtttacggggcccccaggcttcagcaaaagac 1140
Db 1081 gtggtgaaggatgagggctcgatatttctggtttacggggcccccaggcttcagcaaaagac 1140
QY 1141 aacgagtcctcatcagccgggcccaggttcctgcccgggtgcgggttccccag 1200
Db 1141 aacgagtcctcatcagccgggcccaggttcctgcccgggtgcgggttccccag 1200
QY 1201 gtaagtacctggcagccgaggtgtggtcctgcattacacagactggctgcatccccag 1260
Db 1201 gtaagtacctggcagccgaggtgtggtcctgcattacacagactggctgcatccccag 1260
QY 1261 gaccgggcagccctgagggagggccctgagcgatgtggtgggcaccacaatgtcgtgtgc 1320
Db 1261 gaccgggcagccctgagggagggccctgagcgatgtggtgggcaccacaatgtcgtgtgc 1320
QY 1321 cccgtggcccagctggctgggcgactggctgcccaggggtcccgggtctacgcctacgtc 1380
Db 1321 cccgtggcccagctggctgggcgactggctgcccaggggtcccgggtctacgcctacgtc 1380
QY 1381 ttgtgaacacgctgttccacgctcctcctgcccctgtggatgggggtgccccacgggtac 1440
Db 1381 ttgtgaacacgctgttccacgctcctcctgcccctgtggatgggggtgccccacgggtac 1440
QY 1441 gagatcaggttcattctttgggatccccctggacccctctcgaaactacacggcagaggag 1500
Db 1441 gagatcaggttcattctttgggatccccctggacccctctcgaaactacacggcagaggag 1500
QY 1501 aaaatcttgcggccagcactgatggaatactgggccaaactttgccgcacaggggatccc 1560
Db 1501 aaaatcttgcggccagcactgatggaatactgggccaaactttgccgcacaggggatccc 1560
QY 1561 aatgagccccgagacccccaaaggccccacaatggcccccggtacacggcggggctcagcag 1620

Db 1561 aatgagccccgagacccccaaaggccccacaatggccccgtacacggcgggggtcagcag 1620
QY 1621 tacgttagtcttgacctgcggccgctgaggtgcggggggtgcgcgcccaggcctgc 1680
Db 1621 tacgttagtcttgacctgcggccgctgaggtgcggggggtgcgcgcccaggcctgc 1680
QY 1681 gccttctggaaccgcttctcctcccaaatgtcagcgccacc 1722
Db 1681 gccttctggaaccgcttctcctcccaaatgtcagcgccacc 1722

RESULT 3
AAQ99002
ID AAQ99002 standard; DNA; 2256 BP.
XX
AC AAQ99002;
XX
DT 31-MAR-1996 (first entry)
XX
DE Human acetylcholinesterase (AChE) gene.
XX
KW Acetylcholinesterase; acetyl cholinesterase; EC-3.1.1.7;
KW chromosome-7q22; acetylcholine-hydrolyzing enzyme; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 160..2206
FT /*tag= a
XX
PN WO9523158-A1.
XX
PD 31-AUG-1995.
XX
PF 28-FEB-1995; 95WO-US02806.
XX
PR 09-JAN-1995; 95US-0370156.
PR 28-FEB-1994; 94US-0202755.
XX
PA (KOHN/) KOHN K I.
PA (YISS) YISSUM RES & DEV CO.
XX
PI Shani M, Soreq H, Zakut H;
XX
DR WPI; 1995-311499/40.
DR P-PSDB; AAR80726.
XX
PS Claim 3; Fig.1A; 55pp; English.
XX
CC This DNA sequence encoding human acetylcholinesterase is useful
CC for producing transgenic animals which express AChE. The
CC transgenic animals are in turn useful as an assay system for
CC determining the anti-ChE activity of organophosphates, carbamates,
CC anti-ChE drugs, plant glycoalkaloids and snake venoms. This gene
CC contains a promoter region, 6 exons (E1 to E6) and 4 introns (I1-
CC I4).
XX
SQ Sequence 2256 BP; 390 A; 757 C; 680 G; 429 T; 0 other;

Query Match 99.8%; Score 1722; DB 16; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctgtgcacacgccttccctgggttccccactccttcctc 60
Db 160 atgagggccccgcagtgctgtgcacacgccttccctgggttccccactccttcctc 219

Qy 61 ctccctctggctcctgggtggagagtggtgggctgagggccggaggatgcagagctgctg 120
|||||
Db 220 ctccctctggctcctgggtggagagtggtgggctgagggccggaggatgcagagctgctg 279
|||||
Qy 121 gtgacgggtgcgtggggccggctgcggggcattcgcctgaagaccccccgggccctgtc 180
|||||
Db 280 gtgacgggtgcgtggggccggctgcggggcattcgcctgaagaccccccgggccctgtc 339
|||||
Qy 181 tctgctttcctgggcataccccctttgcggagccaccccatgggacccccgcgtctcttgcca 240
|||||
Db 340 tctgctttcctgggcataccccctttgcggagccaccccatgggacccccgcgtctcttgcca 399
|||||
Qy 241 cggagagcccaagcagcccttggtcaggggtggtagacgctacaaccttccagagtgtctgc 300
|||||
Db 400 ccggagcccaagcagcccttggtcaggggtggtagacgctacaaccttccagagtgtctgc 459
|||||
Qy 301 taccaatatgtggacaccctatacccagggttttgaggggcacgagatgtggaaccccaac 360
|||||
Db 460 taccaatatgtggacaccctatacccagggttttgaggggcacgagatgtggaaccccaac 519
|||||
Qy 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccggcctaca 420
|||||
Db 520 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccggcctaca 579
|||||
Qy 421 tcccccacccctgtccctcgtctggaatctatgggggtggctttctacagtggggcctctcc 480
|||||
Db 580 tcccccacccctgtccctcgtctggaatctatgggggtggctttctacagtggggcctctcc 639
|||||
Qy 481 ttggacgtgtacgatggccgcttcttggtacagggccgagaggaactgtgctggtgtccatg 540
|||||
Db 640 ttggacgtgtacgatggccgcttcttggtacagggccgagaggaactgtgctggtgtccatg 699
|||||
Qy 541 aactacccgggtgggagccctttggctcctggccctgcggggagccgagagccccgggc 600
|||||
Db 700 aactacccgggtgggagccctttggctcctggccctgcggggagccgagagccccgggc 759
|||||
Qy 601 aatgtgggtctcctggatcagaggctggccctgcagtggggtgcaggagaacgtggcagcc 660
|||||
Db 760 aatgtgggtctcctggatcagaggctggccctgcagtggggtgcaggagaacgtggcagcc 819
|||||
Qy 661 ttcgggggtgacccgacatcagtgcacgtgtttggggagagcgcggtgagccgctcggtg 720
|||||
Db 820 ttcgggggtgacccgacatcagtgcacgtgtttggggagagcgcggtgagccgctcggtg 879
|||||
Qy 721 ggcattgcacctgtgtccccgcgccagccgggacctgttccacagggccgtgtgcagagc 780
|||||
Db 880 ggcattgcacctgtgtccccgcgccagccgggacctgttccacagggccgtgtgcagagc 939
|||||
Qy 781 ggtgcccccaaatggaccctgggccaagctgggcatgggagagggcccgctgcagggccacg 840
|||||
Db 940 ggtgcccccaaatggaccctgggccaagctgggcatgggagagggcccgctgcagggccacg 999
|||||
Qy 841 cagctggccccacctgtgtgggctgtcctccagggcgcaactggtgggaatgacacagagctg 900
|||||
Db 1000 cagctggccccacctgtgtgggctgtcctccagggcgcaactggtgggaatgacacagagctg 1059
|||||
Qy 901 gtagcctgccttcggacacgacacgagcgaggtcctggtgaaccacgaatggcacgtgctg 960
|||||
Db 1060 gtagcctgccttcggacacgacacgagcgaggtcctggtgaaccacgaatggcacgtgctg 1119
|||||
Qy 961 cctcaagaaaagcgtcttccgggttctccttcgtgcctgtggttagatggagacttccctcagt 1020
|||||
Db 1120 cctcaagaaaagcgtcttccgggttctccttcgtgcctgtggttagatggagacttccctcagt 1179
|||||
Qy 1021 gacacccccagagggccctcatcaacgcggggagacattccacggcctgcaggtgctggggt 1080
|||||
Db 1180 gacacccccagagggccctcatcaacgcggggagacattccacggcctgcaggtgctggggt 1239
|||||
Qy 1081 gtggtgaaggatgaggggtcgtattttctgttttacggggccccaggcttcagcaaaagac 1140
|||||
Db 1240 gtggtgaaggatgaggggtcgtattttctgttttacggggccccaggcttcagcaaaagac 1299
|||||
Qy 1141 aacgagttctctcatcagccgggcccaggttccctggccggggtgcgggtcggggttccccag 1200
|||||

Db 1300 aacgagttctctcatcagccgggcccagagttcctgcccgggtgcgggttcctccag 1359
|||||
Qy 1201 gtaagtacactggcagccgaggtgtgtgctcctgcattacacagactggctgcaccccgag 1260
|||||
Db 1360 gtaagtacactggcagccgaggtgtgtgctcctgcattacacagactggctgcaccccgag 1419
|||||
Qy 1261 gaccggcacgcctgagggagggccctgagcgtgtggtgggagccacaatgtcgtgtgc 1320
|||||
Db 1420 gaccggcacgcctgagggagggccctgagcgtgtggtgggagccacaatgtcgtgtgc 1479
|||||
Qy 1321 cccgtggccccagctggctgggagactggctggccccaggggtgccggggtctacgcctacgtc 1380
|||||
Db 1480 cccgtggccccagctggctgggagactggctggccccaggggtgccggggtctacgcctacgtc 1539
|||||
Qy 1381 tttgaacacccgtgcttccacgctctcctggccctgtggtggtgggggtgccccacggctac 1440
|||||
Db 1540 tttgaacacccgtgcttccacgctctcctggccctgtggtggtgggggtgccccacggctac 1599
|||||
Qy 1441 gagatcgagttcatctttggggtacccccctggacccctctcgaactacacggcagaggag 1500
|||||
Db 1600 gagatcgagttcatctttggggtacccccctggacccctctcgaactacacggcagaggag 1659
|||||
Qy 1501 aaaaattctgcgccagcagctgagcgatactgggccaactttgcgcgcacaggggatccc 1560
|||||
Db 1660 aaaaattctgcgccagcagctgagcgatactgggccaactttgcgcgcacaggggatccc 1719
|||||
Qy 1561 aatgagccccgagacccccaaaggccccacaatggccccgtacacggcggggctcagcag 1620
|||||
Db 1720 aatgagccccgagacccccaaaggccccacaatggccccgtacacggcggggctcagcag 1779
|||||
Qy 1621 tacgttagtcttgacactgcggccgctggtgaggtgcggcggggtgcgcgccagggcctgc 1680
|||||
Db 1780 tacgttagtcttgacactgcggccgctggtgaggtgcggcggggtgcgcgccagggcctgc 1839
|||||
Qy 1681 gccttctggaaccgcttccctcccaaaattgctcagcgccacc 1722
|||||
Db 1840 gccttctggaaccgcttccctcccaaaattgctcagcgccacc 1881
|||||
RESULT 4
AAS17547
ID AAS17547 standard; DNA; 5767 BP.
XX
AC AAS17547;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM034.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;

XX WPI; 2002-055120/07.
DR
XX Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 28-31; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pFM034, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;

Query Match 99.2%; Score 1712; DB 24; Length 5767;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cccccgcagtgtctgtgcacacgccttcctccctggtcccccactccttcctcctcctc 66
DB 844 cccccgcagtgtctgtgcacacgccttcctccctggtcccccactccttcctcctcctc 903

QY 67 tggctcctggtgagagagtggtgggctgagggccgggagagtgatgcagagctgctgtgacg 126
DB 904 tggctcctggtgagagagtggtgggctgagggccgggagagtgatgcagagctgctgtgacg 963

QY 127 gtgcgtggggcggtgcggggcattgcctgaagaccacccatgggaccccggtcctgtctgtct 186
DB 964 gtgcgtggggcggtgcggggcattgcctgaagaccaccccggtcctgtctgtctgtctgt 1023

QY 187 ttccctgggcatcccttttgcggagccaccatgggaccccgctgcctttctgcccacggag 246
DB 1024 ttccctgggcatcccttttgcggagccaccatgggaccccgctgcctttctgcccacggag 1083

QY 247 cccaagcagccttggtcaggggtggtagacgctacaaccttcacagatgtgtgacctaccaa 306
DB 1084 cccaagcagccttggtcaggggtggtagacgctacaaccttcacagatgtgtgacctaccaa 1143

QY 307 tatgtggacacctatataccagggtttttaggggcaccgagatgtggaaccccccaaccgtgag 366
DB 1144 tatgtggacacctatataccagggtttttaggggcaccgagatgtggaaccccccaaccgtgag 1203

QY 367 ctgagcagaggactgcctgtacctcaacgtgtggacaccatacccccgccctacatcccc 426
DB 1204 ctgagcagaggactgcctgtacctcaacgtgtggacaccatacccccgccctacatcccc 1263

QY 427 acccctgtcctcgtctggatctatgggggtggccttctacagtggggcctcctccttgagac 486
DB 1264 acccctgtcctcgtctggatctatgggggtggccttctacagtggggcctcctccttgagac 1323

QY 487 gtgtacgatggccgcttcttgggtacaggccgagagagactgtgtggtgtccatgaactac 546
DB 1324 gtgtacgatggccgcttcttgggtacaggccgagagagactgtgtggtgtccatgaactac 1383

QY 547 cgggtgggagccttggcttccctgcccctgcggggagccgagagcccgggcaatgtg 606
DB 1384 cgggtgggagccttggcttccctgcccctgcggggagccgagagcccgggcaatgtg 1443

QY 607 ggtctcctggatcagaggtggccctgcagtggtgaggaacgtggcagccttcggg 666
DB 1444 ggtctcctggatcagaggtggccctgcagtggtgaggaacgtggcagccttcggg 1503

QY 667 ggtgacccgacatcagtgacgtgttttgggagagcgcgggagccgcctcgttgggcatg 726
DB 1504 ggtgacccgacatcagtgacgtgttttgggagagcgcgggagccgcctcgttgggcatg 1563

QY 727 cacctgctgtccccgcccagccgggctgttccacaggccgctgctcagagcgtgccc 786
DB 1564 cacctgctgtccccgcccagccgggctgttccacaggccgctgctcagagcgtgccc 1623

QY 787 cccaatggacctgggcccacggtgggcatgggagagggcccctgcagggccacgcagctg 846
DB 1624 cccaatggacctgggcccacggtgggcatgggagagggcccctgcagggccacgcagctg 1683

QY 847 gccacaccttgggctgtcctccaggcggcaactggtgaaccacgaatggcagctgtgcctcaa 906
DB 1684 gccacaccttgggctgtcctccaggcggcaactggtgaatgaacagagctggtagccc 1743

QY 907 tgccttcggacacgacacgagcgcaggtcctcctgctgctggtgtagatggagacttccctcagtgacacc 966
DB 1744 tgccttcggacacgacacgagcgcaggtcctcctggtgaaccacgaatggcaactgctgcctcaa 1803

QY 967 gaaagcgtcttcgggttctccttcgtgcctgtggtgtagatggagacttccctcagtgacacc 1026
DB 1804 gaaagcgtcttcgggttctccttcgtgcctgtggtgtagatggagacttccctcagtgacacc 1863

QY 1027 ccagaggccctcatcaaacgcgggagagacttccacagccctgcaggtgcgtggtggtg 1086
DB 1864 ccagaggccctcatcaaacgcgggagagacttccacagccctgcaggtgcgtggtggtg 1923

QY 1087 aaggatgaggctcgtattttctgtggttaacggggcccccaggcttccagcaaacacgag 1146
DB 1924 aaggatgaggctcgtattttctgtggtttacggggcccccaggcttccagcaaacacgag 1983

QY 1147 tctctcatcagccgggcccaggttccctgcctggtggcgaccaccaatgtcgtgtgccccgtg 1206
DB 1984 tctctcatcagccgggcccaggttccctgcctggtggcgaccaccaatgtcgtgtgccccgtg 2043

QY 1207 gacctggcagccgaggtgtggtcctgcattacacagactggctgcctcccgagaccgg 1266
DB 2044 gacctggcagccgaggtgtggtcctgcattacacagactggctgcctcccgagaccgg 2103

QY 1267 gcacgcctgagggagccctgagcgtgtggtggcgaccaccaatgtcgtgtgccccgtg 1326
DB 2104 gcacgcctgagggagccctgagcgtgtggtggcgaccaccaatgtcgtgtgccccgtg 2163

QY 1327 gcccagctggtggcgactggctgcccagggtgcccgggtctacgcctacgtctttgaa 1386
DB 2164 gcccagctggtggcgactggctgcccagggtgcccgggtctacgcctacgtctttgaa 2223

QY 1387 caccgtgttccacgctcctcctggccctgtggtggtgccccagggctacgagatc 1446
DB 2224 caccgtgttccacgctcctcctggccctgtggtggtgccccagggctacgagatc 2283

QY 1447 gagtctatcttgggatccccctggaccctctcgaaactacacggcagagagaaaaac 1506
DB 2284 gagtctatcttgggatccccctggaccctctcgaaactacacggcagagagaaaaac 2343

QY 1507 ttgcggcagcagctgatgcgatactgggccaacttgcggcgacaggggatcccaatgag 1566
DB 2344 ttgcggcagcagctgatgcgatactgggccaacttgcggcgacaggggatcccaatgag 2403

QY 1567 ccccgagaccccccaaggcccccaaatggcccccggtacacggcgggggtcagcagtcagctt 1626

|||||
Db 2404 cccgagacccaaggccccacaatggccccgtacacgcgggggtcagcagtaagtt 2463
QY 1627 agtctggacctgcggccgctggaggtgcggcgggggtgcgcgccccagcctgcgccttc 1686
Db 2464 agtctggacctgcggccgctggaggtgcggcgggggtgcgcgccccagcctgcgccttc 2523
QY 1687 tggaaaccgtctctccccaaattgctcagcgc 1718
Db 2524 tggaaaccgtctctccccaaattgctcagcgc 2555
RESULT 5
AAS17548
ID AAS17548 standard; DNA; 14446 BP.
XX
AC AAS17548;
XX
DT 25-FEB-2002 (first entry)
DE Plasmid pTM036.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigamine bromide; pTM036;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, Soreq H, Arntzen C, Mason H;
XX
DR WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 32-41; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigamine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable

CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM036, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;

Query Match 99.2%; Score 1712; DB 24; Length 14446;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1712; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ccccgagtgctgctgcacagccttcctccctggcttccccactcttctctctctc 66
Db 148 ccccgagtgctgctgcacagccttcctccctggcttccccactcttctctctc 207
QY 67 tggctcctgggtggaggagtgggggctgagggccgggagatgcagagctgctggtgacg 126
Db 208 tggctcctgggtggaggagtgggggctgagggccgggagatgcagagctgctggtgacg 267
QY 127 gtgctggggccggtgcgggcatcctgcctgaagaccccgggggccctgtctctgct 186
Db 268 gtgctggggccggtgcgggcatcctgcctgaagaccccgggggccctgtctctgct 327
QY 187 ttctgggcatccccctttgcggagccaccatgggaccccgctcttctgccaccggag 246
Db 328 ttctgggcatccccctttgcggagccaccatgggaccccgctcttctgccaccggag 387
QY 247 ccaaagcagccttggtcaggggtggtagacgctacaacctccagagtgctgtctaccaa 306
Db 388 ccaaagcagccttggtcaggggtggtagacgctacaacctccagagtgctgtctaccaa 447
QY 307 tatgtggacacccctataccaggttttgagggcacccgagatggtgaaaccccaaccgtgag 366
Db 448 tatgtggacacccctataccaggttttgagggcacccgagatggtgaaaccccaaccgtgag 507
QY 367 ctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccgccctacatcccc 426
Db 508 ctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccgccctacatcccc 567
QY 427 accctgtcctcgtctggtgatctatgggggtgtgcttctacagtggggcctcctctggac 486
Db 568 accctgtcctcgtctggtgatctatgggggtgtgcttctacagtggggcctcctctggac 627
QY 487 gtgtacgatggccgcttctgtgtacagccgagaggactgtgctgttccatgaactac 546
Db 628 gtgtacgatggccgcttctgtgtacagccgagaggactgtgctgttccatgaactac 687
QY 547 cgggtgggagccttttgcttctcctggccctgcggggagccgagagggcccgggcaatgtg 606
Db 688 cgggtgggagccttttgcttctcctggccctgcggggagccgagagggcccgggcaatgtg 747
QY 607 ggtctcctggatcagaggctggccctgcagtggtgcaggagaaacgtggcagccttcggg 666
Db 748 ggtctcctggatcagaggctggccctgcagtggtgcaggagaaacgtggcagccttcggg 807
QY 667 ggtgacccgacatcagtgacgtgtttggggagagcgcgggagcccgctcggtgggcatg 726
Db 808 ggtgacccgacatcagtgacgtgtttggggagagcgcgggagcccgctcggtgggcatg 867
QY 727 cactgtctgtccccgccccagcgggcctgtttccacagggccgctgctgcagagcggtgcc 786
Db 868 cactgtgtgtccccgccccagcgggcctgtttccacagggccgctgctgcagagcggtgcc 927
QY 787 ccaatggaccctggggccacggtgggcatgggagagggccgctgctgcagagcggatg 846
Db 928 ccaatggaccctggggccacggtgggcatgggagagggccgctgctgcagagcggatg 987
QY 847 gccaccttgtgggtgtctccagcgggcactggtgggaatgacacagagctggttagcc 906
Db 988 gccaccttgtgggtgtctccagcgggcactggtgggaatgacacagagctggttagcc 1047

QY 765 ggcgtgctgcagagcggtgcccccaatggaccctgggcccacgtgggcatgggagagggc 824
|||||
Db 921 ggcgtgctgcagagcggtgcccccaatggaccctgggcccacgtggccatgggagagggc 980
QY 825 ccgtgcgagggccacgcagctggcccaccttgtgggctgtcctccaggcggcactggtgg 884
|||||
Db 981 ccgtgcgagggccacccagctggcccaccttgtgggctgtcctccaggcggcactggtgg 1040
QY 885 gaatgacacagagctggtagcctgccttcgacacagacacagcaggtcctggtgaacca 944
|||||
Db 1041 gaatgacacagagctggtagcctgccttcgacacagacacagcaggtcctggtgaacca 1100
QY 945 cgaatggcacgtgctgcctcaagaaagcgtcttcggttctccttcgtgcctggtgtaga 1004
|||||
Db 1101 cgaatggcacgtgctgcctcaagaaagcgtcttcggttctccttcgtgcctggtgtaga 1160
QY 1005 tggagacttctcagtgacacccacagagccctcatcaacggggagagacttccacggcct 1064
|||||
Db 1161 tggagacttctcagtgacacccacagagccctcatcaacggggagagacttccacggcct 1220
QY 1065 gcaagtgctggtggtggtgaaggatgagggtcgattttctggtttacggggcccc 1124
|||||
Db 1221 gcaagtgctggtggtggtgaaggatgagggtcgattttctggtttacggggcccc 1280
QY 1125 aggttccagcaaaagacaaacagtgctctcatcagccgggcccaggttccctggcgggtgcg 1184
|||||
Db 1281 aggttccagcaaaagacaaacagtgctctcatcagccgggcccaggttccctggcgggtgcg 1340
QY 1185 ggtcgggggttccccaggttaagtacacctggcagccgaggtgtggtcctgcattacacaga 1244
|||||
Db 1341 ggtcgggggttccccaggttaagtacacctggcagccgaggtgtggtcctgcattacacaga 1400
QY 1245 ctggtgcattccccagggaccgcgcacctgagggagggccctgagcgtggtggtggcga 1304
|||||
Db 1401 ctggtgcattccccagggaccgcgcacctgagggagggccctgagcgtggtggtggcga 1460
QY 1305 ccacaatgtcgtgtgcccccggtgcccagctggctggcgactggctgcccaggggtgccc 1364
|||||
Db 1461 ccacaatgtcgtgtgcccccggtgcccagctggctggcgactggctgcccaggggtgccc 1520
QY 1365 ggtctacgcctacgtctttgaacacaccgtgcttccacgctctcctggccccctgtggtggg 1424
|||||
Db 1521 ggtctacgcctacgtctttgaacacaccgtgcttccacgctctcctggccccctgtggtggg 1580
QY 1425 ggtgccccacggctacgagatcgagttcatctttgggatccccctggacccctctcga 1484
|||||
Db 1581 ggtgccccacggctacgagatcgagttcatctttgggatccccctggacccctctcga 1640
QY 1485 ctacacggcagaggagaaaaatctcgcccagcgcactgatgcgatactgggccaaacttgc 1544
|||||
Db 1641 ctacacggcagaggagaaaaatctcgcccagcgcactgatgcgatactgggccaaacttgc 1700
QY 1545 ccgcacaggggatcccaatgagccccgagaccccaaggccccacaatggcccccgtagac 1604
|||||
Db 1701 ccgcacaggggatcccaatgagccccgagaccccaaggccccacaatggcccccgtagac 1760
QY 1605 ggcgggggctcagcagtagcttagtctggacctgcggccgctggaggtgcggggggcct 1664
|||||
Db 1761 ggcgggggctcagcagtagcttagtctggacctgcggccgctggaggtgcggcgggggcct 1820
QY 1665 gcgcggccaggcctgcgccttcttggaaccgcttctcccccaaaattgctcagcgcacc 1722
|||||
Db 1821 gcgcggccaggcctgcgccttcttggaaccgcttctcccccaaaattgctcagcgcacc 1878

RESULT 7
AAS17492
ID AAS17492 standard; DNA; 12113 BP.
XX
AC AAS17492;
XX
DT 26-FEB-2002 (first entry)
XX

DE Human acetylcholinesterase, ACHE, gene.
XX
KW Human; ds; ACHE; acetylcholinesterase; polymorphic variant; haplotyping;
genotyping; neurological disease; Parkinson's disease;
KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace (1950,C)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT replace (2237,T)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4044,T)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT CDS 4095..8113
FT /*tag= d
FT /product= "ACHE"
FT exon 4095..5162
FT /*tag= e
FT /number= 2
FT 5163..5508
FT /*tag= f
FT /number= 2
FT replace (4130,T)
FT /*tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4195,A)
FT /*tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4277,C)
FT /*tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4932,G)
FT /*tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4967,T)
FT /*tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT replace (5126,A)
FT /*tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT replace (5151,A)
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT exon 5509..5993
FT /*tag= n
FT /number= 3
FT 5994..6988
FT /*tag= o
FT /number= 3
FT replace (5871,T)
FT /*tag= p
FT /standard_name= "Single nucleotide polymorphism"
FT replace (6831,A)
FT /*tag= q
FT /standard_name= "Single nucleotide polymorphism"
FT exon 6989..7158
FT /*tag= r
FT /number= 4
FT 7159..7991
FT /*tag= s
FT /number= 4
FT 7992..8113
FT /*tag= t
FT /number= 5
FT replace (8114,A)
FT /*tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT replace (8200,A)
FT variation

FT /*tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT replace (8227,A)
FT /*tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT replace (8425,G)
FT /*tag= x
FT /standard_name= "Single nucleotide polymorphism"
XX

PN WO200179219-A2.

XX

PN 25-OCT-2001.

XX

PF 11-APR-2001; 2001WO-US11853.

XX

PR 14-APR-2000; 2000US-197173P.

XX

PA (GENA-) GENAISSANCE PHARM INC.

PA (KAZE/) KAZEMI A.

XX

PI Bentivegna SC, Chew A, Choi JY, Koshy B;

XX

DR WPI; 2002-055248/07.

DR P-PSDB; AAU11231.

XX

CC New polymorphic variants comprising acetylcholinesterase (ACHE)

PT isogene, useful in expressing ACHE protein for use in screening for

PT candidate drugs to treat diseases related to ACHE activity, e.g.

PT neurological diseases or cancer -

XX

PS Claim 21; Fig 1; 79pp; English.

XX

CC The invention relates to a polynucleotide comprising a polymorphic

CC variant of an acetylcholinesterase (ACHE) gene or fragment, protein or

CC complement, the variant comprising an ACHE isogene defined by a haplotype

CC selected from haplotypes 1-20 listed in the specification. Also included

CC are methods for haplotyping and genotyping the ACHE gene of an

CC individual, a method for predicting a haplotype pair for the ACHE gene of

CC an individual, a method for identifying an association between a trait

CC and at least one haplotype or haplotype pair of ACHE gene, recombinant

CC nonhuman organisms transformed or transfected with the polynucleotide

CC where the organism expresses ACHE protein encoded by the first

CC nucleotide sequence or encoded by the polymorphic variant sequence,

CC an isolated antibody specific for and immunoreactive with ACHE,

CC a method of screening for drugs targeting the polypeptide contacting ACHE

CC polymorphic variant with a candidate agent and assaying for binding

CC activity, a computer system for storing and analysing polymorphism data
CC for ACHE gene and a genome anthology for ACHE gene which comprises ACHE
CC isogenes defined by haplotypes 1-20 given in the specification.
CC The Polymorphisms are useful for studying the biological function of
CC ACHE as well as in identifying drugs targeting this protein for the
CC treatment of disorder related to its abnormal expression or function.
CC The polymorphic variants may also be used in screening for compounds
CC targeting ACHE to treat a specific condition or disease predicted to be
CC associated with ACHE activity e.g. neurological diseases (e.g.
CC Parkinson's disease and Alzheimer's disease), cancer, leukaemia,
CC and tumours. The ACHE gene maps to human chromosome 7q22. The
CC present sequence is the ACHE gene.
XX
SQ Sequence 12113 BP; 2215 A; 3836 C; 3766 G; 2296 T; 0 other;

Query Match 62.0%; Score 1070; DB 24; Length 12113;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgagggccccgcagtgctgtctgtgcacacgccttccttgcttccccactctctcctc 60
|||||
Db 4095 atgagggccccgcagtgctgtctgtgcacacgccttccttgcttccccactctcctc 4154

Qy 61 ctctctgtgctcctggtggaggagtggtggggctgagggccgggaggtgcagagctgctg 120
|||||
Db 4155 ctctctgtgctcctggtggaggagtggtggggctgagggccgggaggtgcagagctgctg 4214

Qy 121 gtgacggtgcgtggggccgggtgcggggccttcctgcagaccccccgggggccccctgtc 180
|||||
Db 4215 gtgacggtgcgtggggccgggtgcggggccttcctgcagaccccccgggggccccctgtc 4274

Qy 181 tctgctttcctggggcatcccctttgcgggagccaccctcatgggaccccgctgcttctgcc 240
|||||
Db 4275 tctgctttcctggggcatcccctttgcgggagccaccctcatgggaccccgctgcttctgcc 4334

Qy 241 ccggagccccaagcagccttggtcaggggtggtagacgctacaaaccttccagagtgtctgc 300
|||||
Db 4335 ccggagccccaagcagccttggtcaggggtggtagacgctacaaaccttccagagtgtctgc 4394

Qy 301 taccaatatgtggacacccttatataccacaggttttgagggcaccgagatgtggaaccccaac 360
|||||
Db 4395 taccaatatgtggacacccttatataccacaggttttgagggcaccgagatgtggaaccccaac 4454

Qy 361 cgtgagctgagcagaggactgcctgtacctcaactcaactgtggacaccataccccggcctaca 420
|||||
Db 4455 cgtgagctgagcagaggactgcctgtacctcaactcaactgtggacaccataccccggcctaca 4514

Qy 421 tccccacccctgcctcctgtctggtatctatgggggtggtcttctacagtggggcctcctcc 480
|||||
Db 4515 tccccacccctgcctcctgtctggtatctatgggggtggtcttctacagtggggcctcctcc 4574

Qy 481 ttggacgtgtacgatggccgcttcttctgttacaggccgagagactgtgctggtgtccatg 540
|||||
Db 4575 ttggacgtgtacgatggccgcttcttctgttacaggccgagagactgtgctggtgtccatg 4634

Qy 541 aactaccgggtgggagccttttggtctcctgcctgcggggagccgagagggccccgggc 600
|||||
Db 4635 aactaccgggtgggagccttttggtctcctgcggggagccgagagggccccgggc 4694

Qy 601 aatgtgggtctcctggatcagaggctggccctgcagtggtgtcaggagacgtggcagcc 660
|||||
Db 4695 aatgtgggtctcctggatcagaggctggccctgcagtggtgtcaggagacgtggcagcc 4754

Qy 661 ttcgggggtgaccgcacatcagtgacgctgttttggggagagcgcgggagcgcctcggtg 720
|||||
Db 4755 ttcgggggtgaccgcacatcagtgacgctgttttggggagagcgcgggagcgcctcggtg 4814

Qy 721 ggcatgcacctgtctcccgccagcccggtggccacggtggcatgggagggccctgcagagc 780
|||||
Db 4815 ggcatgcacctgtctcccgccagcccggtggccacggtggcatgggagggccctgcagagc 4874

Qy 781 ggtgcccccaatggaccctgggccacggtggcatgggagggccctgcagggccacg 840
|||||
Db 4875 ggtgcccccaatggaccctgggccacggtggcatgggagggccctgcagggccacg 4934

Qy 841 cagctggccccaccttgtggctgtcctccagggcggcactggtgggaatgacacagagctg 900
|||||
Db 4935 cagctggccccaccttgtggctgtcctccagggcggcactggtgggaatgacacagagctg 4994

Qy 901 gtacgctgccttcggacacgaccagcaggtcctggtgaaccacgaatggcacgtgctg 960
|||||
Db 4995 gtacgctgccttcggacacgaccagcaggtcctggtgaaccacgaatggcacgtgctg 5054

Qy 961 cctcaagaaaagcgtcttccggttctccttctcctgctggtgtagatggagacttccctcagt 1020
|||||
Db 5055 cctcaagaaaagcgtcttccggttctccttctcctgctggtgtagatggagacttccctcagt 5114

Qy 1021 gacaccccccagggccccctcatcaacgcgggagacttccacggccctgcaggt 1070
|||||
Db 5115 gacaccccccagggccccctcatcaacgcgggagacttccacggccctgcaggt 5164

RESULT 8
AAQ05999
ID AAQ05999 standard; DNA; 1800 BP.
XX
AC AAQ05999;
XX
DT 16-JAN-1991 (first entry)

XX	DE	Sequence encoding foetal human acetylcholinesterase (hAChE).
XX	KW	Organophosphorous poisoning; OP; cancer; leukaemia;
KW	KW	megakaryocytopoiesis; ovarian cancer; ds.
XX	OS	Homo sapiens.
XX	FH	Key
FT	CDS	Location/Qualifiers
FT		1..1500
FT	3'UTR	/*tag= a
FT		1501..1800
FT		/*tag= b
XX	XX	EP388906-A.
XX	PD	26-SEP-1990.
XX	PF	20-MAR-1990; 90EP-0105274.
XX	XX	21-MAR-1989; 89IL-0089703.
XX	PA	(YISS) YISSUM RES DEV CO.
XX	PI	Soreq H, Zakut H;
XX	DR	WPI; 1990-291865/39.
DR	DR	P-PSDB; AAR06990.
XX	PT	Human acetylcholinesterase DNA and prodn. of recombinant hAChE -
PT	PT	for treatment of organo-phosphorous poisoning and diagnosis of haemo-
PT	XX	cytopoeitic disorders and ovarian carcinomas.
XX	PS	Disclosure; Fig 1c; 47pp; English.
XX	XX	hAChE is useful as an active pharmacological component for the
CC	CC	prophylaxis and treatment of organophosphorous poisoning, and
CC	CC	post-surgical apnea due to succinylcholine administration.
CC	CC	cDNA probe to the sequence may be used in diagnosis of various
CC	CC	leukaemias, abnormal megakaryocytopoiesis and ovarian carcinomas.
CC	CC	Bases given as N are obscured in the patent specification.
XX	SQ	Sequence 1800 BP; 330 A; 602 C; 539 G; 326 T; 3 other;
Query Match 28.3%; Score 489; DB 11; Length 1800;		
Best Local Similarity 99.5%; Pred. No. 5.1e-217;		
Matches 639; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
Qy	523	actgtgctggtgccatgaactaccgggtgggagcctttggcttcctggccctgccgggg 582
Db	178	actgtgctggtgccatgaactaccgggtgggagcctttggcttcctggccctgccgggg 237
Qy	583	agccgagagggcccgggcaatgtgggtcctcgatcacagaggctggccctgcagtgggtg 642
Db	238	agccgagagggcccgggcaatgtgggtcctcgatcacagaggctggccctgcagtgggtg 297
Qy	643	caggagaaagtggcagccttcgggggtgaccccgacatcatcagtgactgtttggggagagc 702
Db	298	caggagaaagtggcagccttcgggggtgaccccgacatcatcagtgactgtttggggagagc 357
Qy	703	gcgggagccgcctcggtgggcacatgcacctgtgtcccgcccgagccgggacctgtccac 762
Db	358	gcgggagccgcctcggtgggcacatgcacctgtgtcccgcccgagccgggacctgtccac 417
Qy	763	agggccgtgtgcagagcgggtgcccccaatggaccctgggccacggtgggcatgggagag 822
Db	418	agggccgtgtgcagagcgggtgcccccaatggaccctgggccacggtgggcatgggagag 477
Qy	823	gcccgtcgaagggccacgcagctggcccaacttgtgggtgtcctccaggcggaactggt 882
Db	478	gcccgtcgaagggccacgcagctggcccaacttgtgggtgtcctccaggcggaactggt 537

Qy	883	gggaatgacacagagctgtagcctgccttcgggacacgaccagcgaggtcctggtgaac 942
Db	538	gggaatgacacagagctgtagcctgccttcgggacacgaccagcgaggtcctggtgaac 597
Qy	943	cacgaatggcacgtgctgcctcaagaaagcgtcttcctccggttctccttcgtgctggtta 1002
Db	598	cacgaatggcacgtgctgcctcaagaaagcgtcttcctccggttctccttcgtgctggtta 657
Qy	1003	gatggagacttctcctcagtgacacccccagagggccctcatcaacgcgggagacttccacggc 1062
Db	658	gatggagacttctcctcagtgacacccccagagggccctcatcaacgcgggagacttccacggc 717
Qy	1063	ctgcaggtgctggtgggtgtggtgaaggatgaggggtcgtattttctggttacgggggcc 1122
Db	718	ctgcaggtgctggtgggtgtggtgaaggatgaggggtcgtattttctggttacgggggcc 777
Qy	1123	ccaggcttcagcaagacaaacgagtctctcatcagccgggcc 1164
Db	778	ccaggcttcagcaagacaaacgagtctctcatcagccgggcc 819
RESULT 9		
AAT34118		
ID	AAT34118 standard; RNA; 35 BP.	
XX		
AC	AAT34118;	
XX	26-FEB-1997 (first entry)	
XX	Acetylcholinesterase primer, 1522+.	
DE		
XX	Antisense oligonucleotide; BCHE; butyrylcholinesterase; BuChE;	
KW	AS-BCHE; primary bone marrow cell culture; erythropoietin; primer;	
KW	cellular differentiation; anti-cholinesterase; ChE; exposure;	
KW	anti-ChE drug therapy; neurodegenerative disease; hematological disease;	
KW	carbamate insecticide; cocaine analogue; ss.	
XX		
OS	Synthetic.	
XX		
PN	WO9621744-A1.	
XX	18-JUL-1996.	
PD		
XX	11-JAN-1996; 96WO-US00322.	
PF		
XX	09-JAN-1995; 95US-0370204.	
PR		
XX	(KOHN/) KOHN K I.	
PA	(YISS) YISSUM RES & DEV CO.	
PA		
XX	Soreq H, Zakut H;	
PI		
XX	WPI; 1996-342296/34.	
DR		
XX	Screening for pre-disposition to anti-cholinesterase exposure - by	
PT	analysing serum levels of butyrylcholinesterase or screening DNA	
PT	for the presence of BuChE allele(s)	
XX	Example 4; Page 63; 73pp; English.	
PS		
XX		
CC	The sequences given in AAT34116-19 are primers which were used to	
CC	amplify the BCHE (butyrylcholinesterase) gene and the AChE (acetyl-	
CC	cholinesterase) gene. These primers were used to investigate	
CC	fetal protection from anti-ChE poisons that may be associated with	
CC	pregnancy. These primers were used to amplify DNA from placental	
CC	tissue. Different gestational age showed marked different intensities	
CC	of signal of each enzyme gene, as did the levels of transcription of	
CC	each of the enzyme genes at a particular stage. Oligonucleotides	
CC	amplified using these primers may be used to identify patients with	
CC	a genetic predisposition for adverse reactions to anti-cholinesterase	
CC	(ChE) exposure from, e.g. anti-ChE drug therapy for neurodegenerative	
CC	diseases or hematological diseases, or exposure to carbamate	
CC	insecticides or cocaine or cocaine analogues.	

XX SQ Sequence 35 BP; 5 A; 12 C; 8 G; 10 T; 0 other;

Query Match
Best Local Similarity 2.0%; Score 35; DB 17; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1363 cgggtctacgcctacgtctttgaacacacgtgcttc 1397
|||||
Db 1 cgggtctacgcctacgtctttgaacacacgtgcttc 35

RESULT 10
ABI99790
ID ABI99790 standard; cDNA; 2089 BP.
XX AC ABI99790;
XX DT 07-MAR-2002 (first entry)
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:890.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX OS Mus musculus.
XX PN WO200188188-A2.
XX PD 22-NOV-2001.
XX PF 18-MAY-2001; 2001WO-JP04192.
XX PR 18-MAY-2000; 2000JP-0145977.
XX PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI; 2002-034733/04.
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX Claim 2; Page 2244-2245; 2690pp; English.
XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX Sequence 2089 BP; 409 A; 638 C; 596 G; 446 T; 0 other;

Query Match
Best Local Similarity 1.9%; Score 33; DB 24; Length 2089;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1216 gccgaggctgtgtcctgcattacacagactgg 1248
|||||

Db 1216 gccgaggctgtgtcctgcattacacagactgg 1248

RESULT 11
AAS41906
ID AAS41906 standard; DNA; 8418 BP.
XX AC AAS41906;
XX DT 17-DEC-2001 (first entry)
XX DE Genomic sequence #222 encoding novel human enzyme polypeptide.
XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ds.
XX OS Homo sapiens.
XX PN WO200155301-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01239.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465566/50.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing,
PT preventing, treating neural, immune system, muscular, reproductive,
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT diseases -
XX
PS Disclosure; SEQ ID No 2032; 1180pp; English.
XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
CC polynucleotides of the invention can also be used in gene therapy.
CC AAS41685-AAS42192 represent DNA sequences encoding for the novel human
CC enzyme polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 8418 BP; 1951 A; 2317 C; 1938 G; 2212 T; 0 other;

Query Match 1.6%; Score 28; DB 22; Length 8418;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 atactgggccaaactttgcccgcacagg 1554
|||||
Db 1989 atactgggccaaactttgcccgcacagg 2016

RESULT 12
AAS74821
ID AAS74821 standard; cDNA; 816 BP.
XX
AC AAS74821;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #10625.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Llu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG10634.
XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
XX Claim 1; SEQ ID No 10625; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 816 BP; 191 A; 238 C; 199 G; 188 T; 0 other;

Query Match 1.6%; Score 27; DB 23; Length 816;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 atactgggccaaacttggccgcacagg 1553
|||||
Db 252 atactgggccaaacttggccgcacagg 278

RESULT 13
AAS74822
ID AAS74822 standard; cDNA; 1532 BP.
XX AAS74822;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #10626.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG10635.
XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
XX Claim 1; SEQ ID No 10626; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 1532 BP; 348 A; 423 C; 441 G; 320 T; 0 other;

Query Match 1.6%; Score 27; DB 23; Length 1532;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1527 atactgggccaaacttggccgcacagg 1553
|||||
Db 1337 atactgggccaaacttggccgcacagg 1363

RESULT 14
AAD09548
ID AAD09548 standard; cDNA; 1538 BP.
XX AAD09548;
XX 10-SEP-2001 (first entry)
XX Human protease protein-11 (PRTS-11) cDNA.
XX Human; protease protein-11; PRTS-11; cytostatic; hypotensive; antiviral;

KW gastrointestinal disorder; anorexia; dysphagia; cardiovascular disorder;
KW atherosclerosis; vasculitis; autoimmune disorder; inflammatory disorder;
KW Alzheimer's disease; cell proliferative disorder; dermatitis; cirrhosis;
KW acquired immune deficiency syndrome; AIDS; neurological disorder; asthma;
KW developmental disorder; epithelial disorder; eczema; dementia; nootropic;
KW neurological disorder; reproductive disorder; infertility; teratogenesis;
KW immunosuppressive; drug screening; actinic keratosis; cardiac; epilepsy;
KW anaemia; antitumour; gene therapy; antibacterial; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 443..676
FT /*tag= a
FT /product= "Human protease protein-11 (PRTS-11)"
XX
PN WO200146443-A2.
XX
PD 28-JUN-2001.
XX
PF 19-DEC-2000; 2000WO-US34811.
XX
PR 23-DEC-1999; 99US-0172055.
PR 21-JAN-2000; 2000US-0177334.
PR 28-JAN-2000; 2000US-0178884.
PR 02-FEB-2000; 2000US-0179903.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yang J, Baughn MR, Burford N, Au-Young J, Lu DAM, Reddy R, Yue H;
PI Nguyen DB, Tang YT, Yao MG, Lal P;
XX
DR WPI; 2001-418080/44.
DR P-PSDB; AAE04884.
XX

AAS74456
ID AAS74456 standard; cDNA; 496 BP.
XX
AC AAS74456;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #10260.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG10269.
XX

Novel human protease proteins (PRTS) useful for diagnosing, treating,
preventing gastrointestinal, cardiovascular, autoimmune/inflammatory,
cell proliferative disorders associated with abnormal expression of
PRTS -
Claim 5; Page 128; 129pp; English.
The present sequence is human protease protein (PRTS-11) cDNA. Human PRTS
and its nucleic acid molecule are useful for the diagnosis, treatment and
prevention of disorders associated with increased or decreased expression
of PRTS. Examples of such disorders include, gastrointestinal disorder
such as anorexia, dysphagia; cardiovascular disorder such as
atherosclerosis, vasculitis; autoimmune/inflammatory disorder such as
acquired immune deficiency syndrome (AIDS), asthma; cell proliferative
disorder such as actinic keratosis, cirrhosis; developmental disorder
such as epilepsy, anaemia; epithelial disorder such as allergic contact
dermatitis, eczema; neurological disorder such as Alzheimer's disease,
dementia and reproductive disorder such as infertility and teratogenesis.
PRTS DNA is useful for creating 'knockin' humanised animals (pigs) or
transgenic animals (mice or rats) to model human disease. PRTS DNA is
also in useful is gene therapy. PRTS and its immunogenic fragments are
useful for screening libraries of compounds in several drug screening
assays. PRTS is useful for analysing the proteome of a tissue or
cell type.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
Claim 1; SEQ ID No 10260; 103pp; English.
The invention relates to isolated polynucleotide (I) and
polypeptide (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human
diagnostic coding sequences of the invention.
Note: The sequence data for this patent did not appear in the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1538 BP; 344 A; 451 C; 372 G; 371 T; 0 other;
Query Match 1.6%; Score 27; DB 22; Length 1538;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 496 BP; 104 A; 161 C; 137 G; 94 T; 0 other;
Query Match 1.4%; Score 25; DB 23; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1527 atactgggccaactttgcccgcacagg 1553
|||||
Db 481 atactgggccaactttgcccgcacagg 507

Qy 369 gagcgaggactgcctgtacctcaac 393
|||||
Db 27 gagcgaggactgcctgtacctcaac 51

Job time: 16866 sec

THIS PAGE BLANK (USPTO)

• • • • •

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 19:59:53 ; Search time 6064.54 Seconds
(without alignments)
3839.079 Million cell updates/sec

Title: US-09-810-861b-5
Perfect score: 1725
Sequence: 1 atgaggccccgcagtgctt.....aattgctcagcgccacctga 1725

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 13736207 seqs, 6748477542 residues

Word size : 0
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1070	62.0		2080	11	BC001541	BC001541 Homo sapi
2	692	40.1		780	10	BI753192	BI753192 603026066
3	603	35.0		752	10	BI917862	BI917862 603183996
4	569	33.0		732	10	BG708331	BG708331 602672268
5	509	29.5		781	10	BG705869	BG705869 602669361
6	505	29.3		931	10	BI488594	BI488594 603021122
7	502	29.1		534	9	AI332425	AI332425 qq25a10.x
8	464	26.9		786	10	BG703349	BG703349 602685108
9	460	26.7		589	9	AI654042	AI654042 ty61a12.x
10	440	25.5		562	10	BE466391	BE466391 hz21b04.x
11	438	25.4		822	9	AI190022	AI190022 qd35e12.x
12	428	24.8		738	10	BI667712	BI667712 6032933067
13	427	24.8		478	9	AI082805	AI082805 ox74d09.x
14	421	24.4		625	10	BI464924	BI464924 603207402
15	417	24.2		702	10	BI915019	BI915019 603177204
16	412	23.9		574	9	AI207928	AI207928 apl4d04.x
17	405	23.5		456	9	AI081892	AI081892 ox77b06.x

C 18	381	22.1	555	9	AI769167	AI769167 wg34f12.x
C 19	372	21.6	577	9	AI802781	AI802781 wf18c08.x
C 20	370	21.4	405	9	AW050645	AW050645 wz19f11.x
C 21	363	21.0	489	9	AI091215	AI091215 oo22c11.x
C 22	363	21.0	507	9	AW196326	AW196326 xm31h05.x
C 23	352	20.4	990	10	BI489087	BI489087 603021122
C 24	332	19.2	689	10	BG395920	BG395920 602458524
C 25	331	19.2	483	9	AI761619	AI761619 wg66f02.x
C 26	324	18.8	379	9	AI081808	AI081808 ox77e10.x
C 27	321	18.6	474	9	AI761618	AI761618 wg66f01.x
C 28	300	17.4	368	9	AW182170	AW182170 xj71d10.x
C 29	299	17.3	728	10	BG707817	BG707817 602671169
C 30	268	15.5	483	9	AI761609	AI761609 wg66e02.x
C 31	249	14.4	462	10	W92595	W92595 ze05a05.r1
C 32	224	13.0	238	10	BM023426	BM023426 le80c06.y
C 33	219	12.7	348	10	H19772	H19772 yn55f10.r1
C 34	219	12.7	378	10	H21027	H21027 yn66g03.r1
C 35	218	12.6	330	9	AW337310	AW337310 xw83e11.x
C 36	216	12.5	514	9	AA032228	AA032228 zf01f07.s
C 37	215	12.5	396	10	H21132	H21132 yn65b07.r1
C 38	213	12.3	586	9	AA446656	AA446656 zw89g10.s
C 39	212	12.3	547	10	BG707892	BG707892 602671083
C 40	208	12.1	380	10	H19954	H19954 yn53c11.r1
C 41	198	11.5	302	9	AA346027	AA346027 EST52141
C 42	189	11.0	366	10	H21093	H21093 yn65b07.s1
C 43	188	10.9	467	10	N98267	N98267 za28g11.r1
C 44	183	10.6	183	10	BM023149	BM023149 le80c06.x
C 45	181	10.5	387	9	AA805231	AA805231 of49g04.s

ALIGNMENTS

RESULT 1

BC001541
LOCUS BC001541 2080 bp mRNA linear HTC 31-JAN-2002
DEFINITION Homo sapiens, clone IMAGE:3453362, mRNA.
ACCESSION BC001541
VERSION BC001541.1 GI:14705895
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2080)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 4 Row: e Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7710156
This clone has the following problem: no polyA-tail.

FEATURES
Location/Qualifiers
1..2080
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:3453362"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT      319 a   692 c   651 g   418 t
ORIGIN

Query Match      62.0%;   Score 1070;   DB 11;   Length 2080;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1070;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy  1 atgaggccccgcagtgctgtgcacacgccttcctcctggcttcaccactccttcctc 60
Db  98 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCACACTCCTTCTCCTC 157
Qy  61 ctccctctggctcctgggtggaggagtgggggctgagggccgggaggatgcagagctgctg 120
Db  158 CTCCTCTGCTCCTGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGATGCAGAGCTGCTG 217
Qy  121 gtgacggtgcgtg9999ccg99ctgc9999cattgc9999cattgc9999c9999ccctgtc 180
Db  218 GTGACGGTGGTGGGGCCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGGCCCTGTC 277
Qy  181 tctgcttctctgggcatcccccttgcggagaccacccatgggacc99cttcttgcca 240
Db  278 TCTGCTTCTCTGGGCATCCCCCTTTGCGAGCCACCCATGGGACCCGCTCGCTTCTTGCCA 337
Qy  241 ccggagcccaagcagccttggtcaggggtggtagacgctacacacttccagagtgctgc 300
Db  338 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACACCTTCCAGAGTGTCTGC 397
Qy  301 taccaatatgtggacacccctataccaggttttgagggcaccgagatgtggaaccccaac 360
Db  398 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 457
Qy  361 cgtgagctgagcagagactgcctgtacctcaacgtgtggacaccataccccccggcctaca 420
Db  458 CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCGCCCTACA 517
Qy  421 tcccccacccctgtcctcgtctggtgatctatg9999gtggtcttctacagtg9999cctctcc 480
Db  518 TCCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 577
Qy  481 ttggacgtgtacgatggccgcttcttggtacagggccgagaggaactgctg99gtccatg 540
Db  578 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGTGTCCATG 637
Qy  541 aactacegggtggagcctttggttcctcctggcctgc99cctgc99ggagc99agccccgggc 600
Db  638 AACTACCGGTGGGAGCCCTTTGGCTTCTCTGGCCCTGCCGGGGAGCCGAGAGCCCGGGC 697
Qy  601 aatgtgggtcctcctggatcagaggctggccctgcagtg99gtgcaggagaaacgtggcagcc 660
Db  698 AATGTGGGTCTCCTGGATCAGAGGCTTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 757
Qy  661 ttcgggggtgaccccgacatcagtgagcgtgtttg99gagagagc99gggagcc99cctcggtg 720
Db  758 TTCCGGGGGTGACCCGACATCAGTGACCGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 817
Qy  721 ggcatgcaacctgctgtcccccgcccaagcc9999cctgttccacagggccgtgctgcagagc 780
Db  818 GGCATGCACCTGCTGTCCCCGCCACCGGGGCCCTGTTCCACAGGGCCGTGCTGCAGAGC 877
Qy  781 ggtgcccccaatggacccttgggccaaggtgggcatgggagagggcc99cgtcgagggccacg 840
Db  878 GGTGCCCCCAATGGACCCCTTGGGCCACGGTGGGCATGGGAGAGGGCCGCTGCGAGGGCCACG 937
Qy  841 cagctggcccaacctgtgggctgtcctccaggcg99cactggtgggaatgacacagagctg 900
Db  938 CAGCTGGCCCAACCTTGTGGGCTGTCTCCAGGCGGCACCTGGTGGGAATGACACAGAGCTG 997
```

```

Qy  901 gtagcctgccttcggacacgaccagcgcaggtcctgtgtgaaccacgaatggcacgtgctg 960
Db  998 GTAGCCTGCCTTCGGACACGACGACGACGACGACGCTCCTGGTGAACCAACGAATGGCAGTGCTG 1057
Qy  961 cctcaagaaagcgtcttcggttctcctcctcctgctgcctgtggtagatggagacttcctcagt 1020
Db  1058 CCTCAAGAAAGCGTCTTCCGGTTCTCCTTCGTCCTGTGGTAGATGGAGACTTCCTCAGT 1117
Qy  1021 gacacccccagagggccctcatcaacgcgggagacttccacggcctgcaggt 1070
Db  1118 GACACCCCAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGT 1167

RESULT 2
BI753192      780 bp      mRNA      linear      EST 25-SEP-2001
LOCUS      603026066F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196256 5',
DEFINITION      mRNA sequence.
ACCESSION      BI753192
VERSION      BI753192.1 GI:15744770
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 780)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11491 row: d column: 17
High quality sequence stop: 780.

FEATURES
Location/Qualifiers
     source              1..780
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5196256"
                        /clone_lib="NIH_MGC_114"
                        /lab_host="DH10B"
                        /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                        Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                        male brains, age range 23-27 yo. Library is oligo-dT
                        primed and directionally cloned (EcoRV site is destroyed
                        upon cloning). Average insert size 1.5 kb, insert size
                        range 1-3 kb. Library is normalized and enriched for
                        full-length clones and was constructed by C. Gruber
                        (Invitrogen). Research Genetics tracking code 019. Note:
                        this is a NIH_MGC Library."
BASE COUNT      111 a   255 c   260 g   154 t
ORIGIN

Query Match      40.1%;   Score 692;   DB 10;   Length 780;
Best Local Similarity 100.0%;   Pred. No. 6.7e-314;
Matches 692;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy  1 atgaggccccgcagtgctgtgcacacgccttcctcctggcttcaccactccttcctc 60
Db  48 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCGCTTCCCTGGCTTCCCCACTCCTTCTCCTC 107
Qy  61 ctccctctggctcctcctgg99gtggaggagtg99gggctgag99cc99gag99atgacagagctg 120
Db  108 CTCCTCTGGCTCTCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 167
```



```
QY 121 gtgacggtgcgtggggccggtgcggggcattccctgaagaccccccggggcctctc 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 GTGACGGTGCgtggggccggtgcggggcattccgctgaagaccccccggggcctctc 227

QY 181 tctgctttcctgggcatcccccttgcggagccaccatgggacccctgcgtttctgcc 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 228 TCTGCTTCTCTGGGcATCCCCCTTGGCGAGCCACCcATGGGACCCCGTCGCTTCTGCCA 287

QY 241 ccggagcccaagcagccttggtcagggggtggtagacgctacacccctccagagtgtctgc 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 288 CCGGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTACACcCTTCCAGAGTGTCTGC 347

QY 301 taccaatatgtggacaccctataccccagggttttgaggggcaccgagatgtggaaccccaac 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 348 TACCAATATGTGGACACCCcTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 407

QY 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 408 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCCTACA 467

QY 421 tccccacccccctgtcctcgtctggtatctatgggggtggtctctacagtgggcctcctcc 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 468 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 527

QY 481 ttggacgtgtacgatggccgcttcttggtacagggccgagagactgtggtgctccatg 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 528 TTGGACGTGTACGATGGCCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 587

QY 541 aactaccgggtgggagcccttggctcctcgtgcctgcggggagccgagagcccccgggc 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 588 AACTACCGGGTGGGAGCCCTTTGGCTTCCTGGCCCTGCCGGGGAGCCGAGAGGCCCGGGGC 647

QY 601 aatgtgggtcctcgtgatcagaggtgcctcgtcagtggtgagggagaaacgtggcagcc 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 648 AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 707

QY 661 ttcgggggtgacccgacatcagtgacgctgtt 692
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 708 TTCGGGGGTGACCCGACATCAGTGACGCTGTT 739
```

```
RESULT 3
BI917862
LOCUS BI917862 752 bp mRNA linear EST 17-OCT-2001
DEFINITION 603183996F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248006 5',
mRNA sequence.
ACCESSION BI917862
VERSION BI917862.1 GI:16199790
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 752)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1625 row: p column: 23
High quality sequence stop: 752.
Location/Qualifiers
FEATURES
SOURCE 1..752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5248006"
```

```
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
```

```
BASE COUNT 106 a 248 c 251 g 147 t
ORIGIN
```

```
Query Match 35.0%; Score 603; DB 10; Length 752;
Best Local Similarity 100.0%; Pred. No. 4.5e-272;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 atgagggcccgagtgctgtgcacagccttccctggcttccccactctctcctc 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 80 ATGAGGCCCGCAGTGCTGTGCTGCACACGCCTTCCCTGGCTTCCCCACTCCTTCTCCTC 139

QY 61 ctccctcgtcctcctgggtggaggagtgggggtgagggccggaggatgcagagctgctg 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 140 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 199

QY 121 gtgacggtgcgtggggccggtgcggggcattcgcctgaagaccccccgggggcctctgc 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 GTGACGGTGCgtggggccggtgcggggcattcgcctgaagaccccccgggggcctctgc 259

QY 181 tctgtttcctgggcatcccccttgcggagccaccccatgggacccctgcgtttctgcc 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 260 TCTGTCTTCTGGGCATCCCCCTTTCGGAGCCACCCATGGGACCCCTCGCTTCTGCCA 319

QY 241 ccggagcccaagcagccttggtcagggggtggtagacgctacacccctccagagtgtctgc 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 320 CCGGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTACGCTACACcCTTCCAGAGTGTCTGC 379

QY 301 taccaatatgtggacacccctataccccagggttttgaggggcaccgagatgtggaaccccaac 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 380 TACCAATATGTGGACACCCcTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 439

QY 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 440 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCCTACA 499

QY 421 tccccacccccctcctcgtctggtatctatgggggtggtctctacagtgggcctcctcc 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 500 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 559

QY 481 ttggacgtgtacgatggccgcttcttggtacagggccgagagactgtggtgctccatg 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 560 TTGGACGTGTACGATGGCCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 619

QY 541 aactaccgggtgggagcccttggctcctcgtgcctgcggggagccgagagcccccgggc 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 620 AACTACCGGGTGGGAGCCCTTTCGGCTTCTCGGCCTGCCGGGAGCCGAGAGGCCCGGGGC 679

QY 601 aat 603
| | |
Db 680 AAT 682
```

```
RESULT 4
BG708331
LOCUS BG708331 732 bp mRNA linear EST 07-MAY-2001
DEFINITION 602672268F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795048 5',
mRNA sequence.
ACCESSION BG708331
VERSION BG708331.1 GI:13985566
KEYWORDS EST.
```

SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 732)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10677 row: g column: 17
High quality sequence stop: 732.

FEATURES
source
1. 732
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4795048"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 101 a 250 c 239 g 142 t
ORIGIN

Query Match 33.0%; Score 569; DB 10; Length 732;
Best Local Similarity 99.8%; Pred. No. 4.1e-256;
Matches 619; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgagccccgcagtgctgtgcacacgccttcctggtccccactctctcctc 60
|||||
Db 113 ATGAGGCCCCGACGTGCTGTGTGCACACGCCCTTCCTGGCTTCCCCACTCTTCTCCTC 172

QY 61 ctccctctggctcctgggtggaggagtgggggctgagggccgggaggatgcagagctgtg 120
|||||
Db 173 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGTGCTG 232

QY 121 gtgacgggtgcgtggggccggctgcgtgcggggcattgcgcctgaagacccccggggccctgtc 180
|||||
Db 233 GTGACGGTGC GTGGGGCCCGGCTGCGGGGCAATCGCCCTGAAGACCCCGGGGGCCCTGTG 292

QY 181 tctgcttctcctgggcatcccccttgcggagccaccatgggacccccgtcttcttgcga 240
|||||
Db 293 TCTGCTTCTCGGCATCCCTTTGCGGAGCCACCCCATGGACCCCGTCGCTTCTGCCA 352

QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttcagagtgctgc 300
|||||
Db 353 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGTGTAGACGCTACAACCTTCCAGAGTGTCTGC 412

QY 301 taccaatatgtggacacccctataccccaggttttgaggggcaccgagatgtggaaccccaac 360
|||||
Db 413 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 472

QY 361 cgtgagctgagcaggagactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
|||||
Db 473 CGTGAGCTGAGCGGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCCTACA 532

QY 421 tccccaccctgcctcgtctggtgatctatgggggtggtcttacagtggggcccctcctcc 480
|||||
Db 533 TCCCCCAGCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCCTCC 592

QY 481 ttggacgtgtacagtgccgcttcttgggtacagtcgagagagactgtgctggtgtccatg 540
|||||
Db 593 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 652

QY 541 aactaccgggtggagccttggcttcctcctggcctgcggggagccgagagagcccgggc 600
|||||
Db 653 AACTACCGGTTGGAGCCTTGGCTTCTCGCCCTGCCGGGAGCCGAGAGGCCCGGGC 712

QY 601 aatgtgggtctcctggatca 620
|||||
Db 713 AATGTGGGTCTCCTGGATCA 732

RESULT 5
BG705869 602669361F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4792046 5',
LOCUS mRNA sequence.
DEFINITION BG705869
ACCESSION BG705869
VERSION BG705869.1 GI:13980645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 781)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10669 row: j column: 15
High quality sequence stop: 781.

FEATURES
source
1. 781
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792046"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 110 a 260 c 262 g 149 t
ORIGIN

Query Match 29.5%; Score 509; DB 10; Length 781;
Best Local Similarity 99.8%; Pred. No. 6.1e-228;
Matches 629; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 atgagggccccgcagtgctgtgcacacgccttcctggcttccccactctctcctc 60
|||||
Db 113 ATGAGGCCCCCGCAGTGCTGTGTGCACACGCCCTTCCCTGGCTTCCCCACTCTTCTCCTC 172

Qy	61	ctcctctggctcctgggtggagagtggtgggctgagggccgggaggatgcagagctgctg	120
Db	173	CTCCTCTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG	232
Qy	121	gtgacggtgcgtggggccggctgcggggcatttcgctgaagaccccccggggcccctgtc	180
Db	233	GTGACGCTGCGTGGGGCCGGCTTCGCGGGCATTCGCCCTCAAGACCCCCCGGGGCCCTGTC	292
Qy	181	tctgccttctcctgggcatacccttttcgcgagcaccacatggggacccccgcctcttctgccca	240
Db	293	TCTGCTTTCCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTCGTGCCA	352
Qy	241	ccggagcccaagcagccttggtcagggggtggtagacgtacacaaccttccagagtctctgc	300
Db	353	CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC	412
Qy	301	taccaatatgtggacacccctataccccaggttttgagggcaccgagatgtggaaacccccaac	360
Db	413	TACCAATATGTGGACACCCCTATACCCAGGTTTGTGAGGGCACCGAGATGTGGAACCCCAAC	472
Qy	361	cgtgagctgagcgaggactgcctgttacctcaacgtgtggacaccatacccccgggcctaca	420
Db	473	CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGCTGTGGACACCATACCCCGGCGCTACA	532
Qy	421	tcccccacccctgtcctcgtctggaatctatgggggtggcttctacagtggggcctcctcc	480
Db	533	TCCCCCA-CCCTGTCTCGTCTGATCTATGGGGTGGCTTCTACAGTGGGGCCCTCCTCC	591
Qy	481	ttggacgtgtacgatggccgcttcttgggtacagggccgagaggactgtgctggtgtcccatg	540
Db	592	TTGGACGCTGTACGATGGCCGCTTCTTGGTACAGGCCCGAGAGGACTGTGCTGGTGTCCATG	651
Qy	541	aactacccgggtgggagacctttagcttccctggccctgcacgggggagccgagagggccccgggc	600
Db	652	AACTACCCGGGTGGAGCCCTTTGGCTTCCTGGCCCTGCCCGGGAGCCGAGAGGCCCGGGGC	711
Qy	601	aatgtgggtctcctggatcagaggctggcc	630
Db	712	AATGTGGGTCTCCTGGATCAGAGGCTGGCC	741

RESULT	6	
BI488594		
LOCUS		931 bp mRNA linear EST 28-AUG-2001
DEFINITION		603021122F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191829 5', mRNA sequence.
ACCESSION		BI488594
VERSION		BI488594.1 GI:15327822
KEYWORDS		EST.
SOURCE		human.
ORGANISM		Homo sapiens
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS		1 (bases 1 to 931)
TITLE		NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT		Unpublished (1999)
		Contact: Robert Strausberg, Ph.D.
		Email: cgapbs-r@mail.nih.gov
		Tissue Procurement: Life Technologies, Inc.
		cDNA Library Preparation: Life Technologies, Inc.
		cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
		DNA Sequencing by: Incyte Genomics, Inc.
		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
		Plate: LLAM11479 ROW: 1 column: 06
		High quality sequence stop: 771.
FEATURES		Location/Qualifiers
source		1. .931
		/organism="Homo sapiens"
		/db xref="taxon:9606"

```

/clone="IMAGE:5191829"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."

```

BASE COUNT	152 a	294 c	306 g	179 t
ORIGIN				
Query Match 29.3%; Score 505; DB 10; Length 931;				
Best Local Similarity 99.8%; Pred. No. 4.8e-226;				
Matches 625; Conservative 0; Mismatches 0; Indels 1; Gaps 1;				
QY	130	cgtggggccgctgcggggcattcgccctgaagaccccccgggggccctgtctctgctttc	189	
DB	1	CGTGGGGCCGGCTCGGGGCATTCGCCTGAAGACCCCCCGGGGGCCCTGCTCTGCTTTC	60	
QY	190	ctgggcaccccccttgggagagcaccacccatgggacccccgtcgctttctgccaccggagccc	249	
DB	61	CTGGGCATCCCCCTTTGGGAGACCACCCATGGGACCCCGTGGCTTCTGCCACCGGAGCCC	120	
QY	250	aagcagccttggtcagggtggtagacgtacaaccttcagagtgctgtactaccaatat	309	
DB	121	AAGCAGCCTTGGTCAGGSGTGGTAGACGCTACAACCTTCCAGAGTGTCTGTACCAATAT	180	
QY	310	gtggacaccctatacccaggttttgaggccacggagatgtggaaccccccaaccgtgagctg	369	
DB	181	GTGGACACCCCTATACCCAGGTTTTGAGGGCACCGAGATGTGGAACCCCAACCGTGAGCTG	240	
QY	370	agcaggactgcctgtacctcaacgtgtggacaccatacccccgccctacatccccacc	429	
DB	241	AGCGAGGACTGCCCTGTACTCAACGTGTGGACACCATACCCCGGGCCTACATCCCCCACC	300	
QY	430	cctgtcctcgtctggatctatgggggtggttctctacagtggggccctcctccttg-gacgt	488	
DB	301	CCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCCTTGTGACGT	360	
QY	489	gtacgatggccgcttcttgggtacagggccgagaggactgtgctggtgtccatgaactaccg	548	
DB	361	GTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATGAACACCG	420	
QY	549	ggtaggagcctttggcttctgctggccctgccggggagccgagagggccccgggcaatgtggg	608	
DB	421	GGTGGAGCCCTTGGCTTCTGGCCCTGCCGGGAGCCGAGAGGCCCGGGCAATGTGGG	480	
QY	609	tctcctggatcagaggtcggccctgcagtggtgtaggagaacgtggcagccttcggggg	668	
DB	481	TCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCCTTCGGGGG	540	
QY	669	tgacccgacatcagtgacgctgttggggagagcgcgggagccgcctcggtgggcatgca	728	
DB	541	TGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGGAGCGCGCGCTCGGTGGGCATGCA	600	
QY	729	cctgctgtccccgccagccggggcc	754	
DB	601	CCTGTCTGTCCCGCCAGCCGGGGCC	626	

RESULT	7
AI332425/c	
LOCUS	
DEFINITION	AI332425 534 bp mRNA linear EST 13-FEB-1999 q25a10.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1933530 3' similar to gb:N55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION	AI332425
VERSION	AI332425.1 GI:4068984

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 534)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 618 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 460.
FEATURES Location/Qualifiers
source
1..534
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1933530"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 95 a 170 c 173 g 95 t 1 others
ORIGIN
Query Match 29.1%; Score 502; DB 9; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.le-224;
Matches 502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1119 ggccccaggcttcagcaaaagacaacgagtcctctcatcagcgggcccaggttcctgcggg 1178
Db 502 GGGCCAGGCTTCAGCAAAAGACAACGAGTCTCTCATCAGCGGGCCGAGTTCCTGGCCGG 443
QY 1179 ggtgcggggtcggggttccccaggtaagtgaagtgcctggcagcggaggtgtggtcctgcatta 1238
Db 442 GGTGCGGGTGGGGTTCCCCAGGTAAGTGACCTGGCAGCCGAGGCTGTGCTGCAATTA 383
QY 1239 cacagactggctgcatacccgaggaccggcacgcctcctgagggagccctgagcgaatgtggt 1298
Db 382 CACAGACTGGCTGCATCCCGAGGACCCGGCACGCCTGAGGAGGCCCTGAGCGATGTGGT 323
QY 1299 gggcgaccacaatgtcgtgtgccccgtgcccagctggctggggcgagctggctgcccaggg 1358
Db 322 GGGCGACCACAATGTGTCGTGTCGCCGTGGCCAGCTGGCTGGCGGACTGGCTGCCCCAGGG 263
QY 1359 tgcgcgggtctacgcctacgtctttgaacacccgtgtcttcacgcctcctcctgccccctgtg 1418
Db 262 TGCCCGGGTCTACGCCCTACGCTTGTGAACACCGTGCTTCACAGCTCTCCTGGCCCCCTGTG 203
QY 1419 gatgggggtgccccacggctacgagatcgagttcatcttttgggatccccctggacccctc 1478
Db 202 GATGGGGGTGCCCCACGGCTACGAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTC 143
QY 1479 tcgaaactacacggcagagagagaaaaatcttcgccacgagcactgatgcgatactggggccaa 1538
Db 142 TCGAAACTACACGGCAGAGGAGAGAAAATCTTCGCCACGCGACTGATGCGATACTGGGCCAA 83

QY 1539 ctttgccgcacaggggatcccaatgagccccgagaccccccaagggccccacaatggcccc 1598
Db 82 CTTTGGCCGCACAGGGGATCCCAATGAGCCCCGAGACCCCAAGGCCCCACAATGGCCCC 23
QY 1599 gtacacggcggggggtcagcag 1620
Db 22 GTACACGGCGGGGGCTCAGCAG 1
RESULT 8
BG703349
LOCUS 602685108F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4817881 5',
DEFINITION mRNA sequence.
ACCESSION BG703349 786 bp mRNA linear EST 07-MAY-2001
VERSION BG703349.1 GI:13975591
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10718 row: o column: 02
High quality sequence stop: 731.
FEATURES Location/Qualifiers
source
1..786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4817881"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: Sali-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 110 a 263 c 261 g 152 t
ORIGIN
Query Match 26.9%; Score 464; DB 10; Length 786;
Best Local Similarity 99.6%; Pred. No. 8.2e-207;
Matches 564; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 atgagggccccgcagtgctgtgcacacgccttcctggctccccactctctcctc 60
Db 113 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCTTCCTGGCTCCCACTCCTCTCCTC 172
QY 61 ctctctggctcctcctgggtgagagtggtgggtgagggccggaggatgcagagctgctg 120
Db 173 CTCCTCTGGCTCCTGGGTGGAGAGTGGGGGTGAGGGCCGGAGGATGCAGAGCTGCTG 232
QY 121 gtgacggtgcgtgggggcccggctgcggggcattgcctgaagacccccggggccctgtc 180
Db 233 GTGACGGTGCCTGGGGGGCCGGCTGCGGGGGCATTCGCCCTGAACACCCCCGGGGCCCTGTC 292

QY	181	tctgctttctgggcataccccctttgcggagccaccccatgggaccccgctcgctttctgcc	240
Db	293	TCTGCTTTCTCTGGGCATCCCCCTTTGCGAGCCACCCATGGGACCCCGTCGCTTTCTGCCA	352
QY	241	ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttcacagagtgtctgc	300
Db	353	CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC	412
QY	301	taccaatatgtggacaccctatacccaggttttgagggcaccgagatgtggaacccaac	360
Db	413	TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC	472
QY	361	cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccgccctaca	420
Db	473	CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCGGGCTACA	532
QY	421	tcccccacccctgtcctcgtctcgtgatctatgggggtggcttctctacagtggggcctctcc	480
Db	533	TCCCCCATCCCTGTCTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC	592
QY	481	ttggacgtgtacgatggccgcttcttggttacagggccgagaggactgctggtgtccatg	540
Db	593	TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGTGTGTCCATG	652
QY	541	aactaccgggtgggagcctttggctt	566
Db	653	AACTACCGGTGGGAGCCTTTGGCTT	678
RESULT	9		
AI654042/c			
LOCUS	AI654042	589 bp	mRNA linear EST 07-MAR-2000
DEFINITION	ty61a12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2283550 3' similar to gb:M55040 ACETYLCOLINESTERASE PRECURSOR (HUMAN);, mRNA sequence.		
ACCESSION	AI654042		
VERSION	AI654042.1	GI:4738021	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 589)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Insert Length: 850 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 461. Location/Qualifiers 1..589 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2283550" /clone_lib="NCI_CGAP_Kid11" /lab_host="DH10B" /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive		
FEATURES	Source		

hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo.			
BASE COUNT	106 a	193 c	186 g 101 t 3 others
ORIGIN			
Query Match	26.7%;	Score 460;	DB 9; Length 589;
Best Local Similarity	99.6%;	Pred. No. 5.9e-205;	
Matches	560;	Conservative 0;	Mismatches 2; Indels 0; Gaps 0;
QY	1059	cggcctgcaggtgctggtggtggtggtgaaggatgaggctcgtattttctggttacgg	1118
Db	562	CGGCCTGCAGGTGCTGGTGGGTGTGGTGAAGGATGAGGGCTCGTATTTTCTGGTTACGG	503
QY	1119	ggccccaggtctcagcaaaagacaacgagtctctcatcagccggccgagttcctggccgg	1178
Db	502	GGCCCCAGGCTTCAGCAAAAGACAACGAGTCTCTCATCAGCCGGCCGAGTCTCTGGCCGG	443
QY	1179	ggtcggggtcggggttccccaggtaaagtacctggcagccgaggtgtggtcctgcatta	1238
Db	442	GGTGGGGTCCGGGTTCCCCAGGTAAGTGACCTGGCAGCCGAGGCTGTGCTCTGCATTA	383
QY	1239	cacagactggctgcatcccgaggacccggcaccgctgaggaggccctgagcagatgtggt	1298
Db	382	CACAGACTGGCTGCATCCCGAGGACCCCGGCACGCCTGAGGGAGGCCCTGAGCGATGTGT	323
QY	1299	ggcgacaccacaatgtcgtgtgcccccgctggccccagctggctggcgactggtgccccagg	1358
Db	322	GGCGGACCCACAATGTCTGTGCNCCGTGGCCACGCTGGCTGGCGACTGGCTGCCACGGG	263
QY	1359	tgccccggtctacgcctacgtcttttgaacacggtgcttccacgctctcctgccccctgtg	1418
Db	262	TGCCCGGGTCTACGCCCTACGTCTTTTGAAACCCCGTGTCTCCACGCTCTCCTGGCCCCCTGTG	203
QY	1419	gatgggggtgccccacggctacgagatcgagttcatctttgggataccccctggacccctc	1478
Db	202	GATGGGGTGCCCCACGGCTACGAGATCGAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTC	143
QY	1479	tcgaaactacggcgagagagagaaaaatcttcgccccagcgactgatcgatactgggcca	1538
Db	142	TCGAAACTACACGGCAGAGGAGAGAAATCTTCGCCACGCGACTGATCGGATACTGGGCCAA	83
QY	1539	ctttgccccgacaggggatcccaatgagccccgagacccccaaagcccccaaatggcccc	1598
Db	82	CTTTGCCCGCACAGGGGATCCCAATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCC	23
QY	1599	gtacacggcgggggtcagcag	1620
Db	22	GTACACGGCGGGGCTCAGCAG	1
RESULT	10		
BE466391/c			
LOCUS	BE466391	562 bp	mRNA linear EST 27-JUL-2000
DEFINITION	hz21b04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3208591 3' similar to gb:M55040 ACETYLCOLINESTERASE PRECURSOR (HUMAN);contains element MSR1 repetitive element ;, mRNA sequence.		
ACCESSION	BE466391		
VERSION	BE466391.1	GI:9512253	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 562)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D.		

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 469.

FEATURES
source
1. .562
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3208591"
/clone_lib="NCI_CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 104 a 184 c 173 g 97 t 4 others
ORIGIN

Query Match 25.5%; Score 440; DB 10; Length 562;
Best Local Similarity 99.8%; Pred. NO. 1.4e-195;
Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1137 agacaacgagtcctcatcagccggccgagtcctctgcccgggtgcgggtcgggttcc 1196
|||||
Db 491 AGACAAACGAGTCTCTCATCAGCCGGCCGAGTTCCTGGCCGGGTCCGGGTTC 432

QY 1197 ccaggtaagtacctggcagcggaggtgtgtctctgcattacacagactggctcatcc 1256
|||||
Db 431 NCAGGTAAGTGACCTGGCAGCCGAGGCTGTGTCTGCATTACACAGACTGGCTCATCC 372

QY 1257 cgaggaccggcacgcctgagggagccctgagcgatgtggtggcgaccacaatgtcgt 1316
|||||
Db 371 CGAGGACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGTGGCGACCAATGTCTGT 312

QY 1317 gtgccccgtggcccagctggctgggcgactggctggccagggtgccccgtctacgccta 1376
|||||
Db 311 GTGCCCCGTGGCCACGCTGGCTGGCGACTGSGTGCCCCAGGGTGCCCCGGTCTACGCCTA 252

QY 1377 cgtctttgaacacacgctgtctccacgctctctcctgcccctgtggtatgggggtgccccacgg 1436
|||||
Db 251 CGTCTTTGAACACCGTGCTTCCACGCTCTCTCTGGCCCTGTGGATGGGGGTGCCCCACGG 192

QY 1437 ctacgagatcgagttcatctcttgggatccccctggaccctctcgaaaactacacggcaga 1496
|||||
Db 191 CTACGAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGA 132

QY 1497 ggagaaaaattctgcccagcagctgatgcgatactgggccaaactttgccgcacacaggga 1556
|||||
Db 131 GGAGAAAATCTTCGCCAGCGACTGATCGGATACTGGGCCAACTTTGCCCGCACAGGGGA 72

QY 1557 tcccaatgagccccgagaccccgaaggccccacaatggccccgtacacggggggtcga 1616
|||||
Db 71 TCCCAATGAGCCCCGAGACCCCAAGGCCCCCAATAATGGCCCCCGTACACGGCGGGGGCTCA 12

QY 1617 gcagtacgtta 1627
|||||
Db 11 GCAGTACGTTA 1

RESULT 11
AI190022/c
LOCUS
DEFINITION
qd35e12.x1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
IMAGE:1731502 3' similar to gb:M5040 ACETYLCOLINESTERASE
PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AI190022
VERSION AI190022.1 GI:3741231
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1019 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 469.
FEATURES
source
1. .822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1731502"
/clone_lib="Soares_fetal_heart_NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGGAGCGCCGCATCTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
BASE COUNT 151 a 271 c 245 g 150 t 5 others
ORIGIN

Query Match 25.4%; Score 438; DB 9; Length 822;
Best Local Similarity 99.6%; Pred. NO. 1.3e-194;
Matches 538; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1106 ttctggtttacggggcccccaggcttcagcaaaagacagagtcctctcatcagcggccg 1165
|||||
Db 540 TTCTGTTTACGGGGCCCCAGGCTTCAGCAAAGACAACGAGTCTCTCATCAGCCGGCCG 481

QY 1166 agttcctggccgggtgcgggtcggggttcccaggttaagtgcacctggcagccagggctg 1225
|||||
Db 480 AGTTCCTGGCCGGGTGCGGGTTCGGGGTTCGCCAGGTAAGTGACCTGGCAGCCGAGGCTG 421

QY 1226 tggctcctgcattacagagactggctgcattcccagagaccggcacgcctgagggagggccc 1285
|||||
Db 420 TGGTCTCGCATTACACAGACTGGCTGCATCCCGAGGACCCCGCACGCTGAGGGAGGCCC 361

QY 1286 tgagcgaatggtggcgaccacaatgtcgtgtgccccgtgccccagctggtggtggcgac 1345
|||||
Db 360 TGAGCGATGTGTTGGCGGACCAACAATGTCTGTTCNCCGTGGCCAGCTGCTGGCGGAC 301

QY 1346 tggctgccccaggggtgccccgggtctacgcctacgtctttgaacaccgtgcttccacgtct 1405
|||||

Db	300	TTGGTGTCCCGAGGGTGCCCGGGTCTACGCCTACGTCTTTGAACACCGTGTCTTCCACGCTCT	241
Qy	1406	cctggccccctgtgatgggggtgccccacggctacgagatcgagttcatcttttggatcc	1465
Db	240	CCTGGCCCCCTGTGGATGGGGGTGCTTCACGGCTACGAGATCGAGTTCATCTTTGGGATCC	181
Qy	1466	ccctggacccctctgaaactacacggcgagaggagaaaaatcttcgccacgagcactgtgc	1525
Db	180	CCCTGGACCCCTCTCGAACTACACGGCAGAGGAGAGAAATCTTCGCCACGCGACTGATGC	121
Qy	1526	gatactggcccaacttgcctgcacaggggatcccaatgagccccgagacccccaaagccc	1585
Db	120	GATACTGGGCCAACTTGCCCGCACAGGGGATCCCAATGAGCCCCGAGACCCCAAGGCC	61
Qy	1586	cacaatggccccctacacggcgggggctcagcagtagcttagtctggacctgcggccgc	1645
Db	60	CACAATGGCCCCCTACACGGCGGGGGCTCAGCAGTACGTTAGTCTGGACCTCGCGCCGC	1
RESULT 12			
BI667712			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Qy			

normalized libraries (melanocyte 2NbHM, pregnant uterus NBHPU, and fetal heart NBHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 84 a 154 c 159 g 81 t
ORIGIN

Query Match 24.8%; Score 427; DB 9; Length 478;
Best Local Similarity 99.8%; Pred. No. 1.8e-189;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1143 cgagtctcatcagcgggccgagttcctgcccggggtgcgggtcggggttccccaggt 1202
|||||
Db 478 CGAGTCTCTCATCAGCGGGCCGAGTTCCTGCGCGGGTGCGGGTCCGCCAGGT 419
QY 1203 aagtgaacctggcagcgaggctgtggtcctgcattacacagactggctgcatccccagga 1262
|||||
Db 418 AAGTGACCTGGCAGCGAGGCTGTGGTCTCGATTACACAGACTGGCTGCATCCCGAGGA 359
QY 1263 cccggcacgcctgagggagggccctgagcgatgtggtgggcgaccacaaatgctgtgtccc 1322
|||||
Db 358 CCCGGCACGCCTGAGGAGGCCCTGAGCGATGTGGTGGCGGACCACAAATGCTGTGTGCC 299
QY 1323 cgtggccagctggctggcgactggctgcccagggtgccccgggtctacgcctactgtctt 1382
|||||
Db 298 CGTGGCCAGCTGGCTGGCGGACTGGCTGCCAGGGTGCCGGGTCTACGCCTACGTCTT 239
QY 1383 tgaacacctgtcttcacgcctcctgcccctctggtgatggggtgccccacagggctacga 1442
|||||
Db 238 TGAACACCGTGTTCACAGCTCTCCTGGCCCTCTGGATGGGGTGCCTCACGGCTACGA 179
QY 1443 gatcaggttcattcttgggatccccctggaccctctcgaactacacggcagaggagaa 1502
|||||
Db 178 GATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAACTACACGSCAGAGGAGAA 119
QY 1503 aatcttcgccagcagctgatcgatactgggccaacttggcccgcacaggggatcccaa 1562
|||||
Db 118 AATCTTCGCCCCAGCGACTGATCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCCAA 59
QY 1563 tgagccccgagagaccccaaggccccacaaatggcccccgctacacggcggtcagcag 1620
|||||
Db 58 TGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1

RESULT 14
BI464924 BI464924 625 bp mRNA linear EST 21-AUG-2001
LOCUS 603207402F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5273017 5',
DEFINITION mRNA sequence.
ACCESSION BI464924
VERSION BI464924.1 GI:15255580
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 625)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11689 row: c column: 02
High quality sequence stop: 625.

FEATURES
source

1. 625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5273017"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 88 a 218 c 196 g 123 t
ORIGIN

Query Match 24.4%; Score 421; DB 10; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.2e-186;
Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaggccccgcagtgctctgctgcacagccttccctggcttccccactctctctc 60
|||||
Db 116 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCGCTTCCCTGGCTTCCCCACTCCTTCTCCTC 175
QY 61 ctctcttgctcctgggtgggagagtggtggggtgagggccgggaggtgcagagctgctg 120
|||||
Db 176 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGCGCGGAGGATGCAGAGCTGCTG 235
QY 121 gtgacggtgcgtgggcccggctgcggggcattgcctgaagacccccccggggccctgtc 180
|||||
Db 236 GTGACGGTGCCTGGGGSCGGCTGCGGGGCATTTCGCTGAAGACCCCCCGGGGGCCTGTC 295
QY 181 tctgcttctcctgggcatcccccttttgcggagccaccatgggacccccgtgcttctctgcc 240
|||||
Db 296 TCTGCTTCTCTGGGCATCCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTCTTGCCA 355
QY 241 ccggagcccaagcagccttggtcaggggtggtgtagcgtacacaccttccagagtgctgtc 300
|||||
Db 356 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACCTTCCAGAGTGTCTGTC 415
QY 301 taccaatatgtggacacctatataccagggttttgggggaccgagatgtggaaccccaac 360
|||||
Db 416 TACCAATATGTGGACACCTATACCCAGGTTTGGGGCACCCGAGATGTGGAACCCCAAC 475
QY 361 cgtgagctgagcagagactgcctgtacctcaacgtggtggacaccatacccccggtctaca 420
|||||
Db 476 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGCTGTGGACACCATACCCCGGCTTACA 535
QY 421 t 421
Db 536 T 536

RESULT 15
BI915019 BI915019 702 bp mRNA linear EST 16-OCT-2001
LOCUS 603177204F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241624 5',
DEFINITION mRNA sequence.
ACCESSION BI915019
VERSION BI915019.1 GI:16179103
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 702)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11609 row: g column: 01
High quality sequence stop: 702.
Location/Qualifiers
1. .702
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5241624"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

FEATURES
source

BASE COUNT 99 a 240 c 226 g 137 t
ORIGIN

Query Match 24.2%; Score 417; DB 10; Length 702;
Best Local Similarity 99.6%; Pred. No. 9.4e-185;
Matches 517; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgagggcccgagtgctgtcgtgcacagcccttccctggcttccccactccttctc 60
Db 87 ATGAGGCCCCCGCAGTGTCTGCTGCACAGCCCTTCCCTGGCTTCCCCACTCCTTCTCCTC 146
Qy 61 ctcctctggctcctcgtggtgagagagtgagggggctgagggccgggagatgcagagctgctg 120
Db 147 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 206
Qy 121 gtgacggtgctggtggggccggtcgcggggcattcgcctgaagaccccgggggccctgtc 180
Db 207 GTGACGGTGCCTGGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGGCCCTGTC 266
Qy 181 tctgcttctcgtgggcaccccttgcggagccacccatgggaccccgctcgttcttgcca 240
Db 267 TCTGCTTTCCTGGGCATCCCTTTGCGAGCCACCCATGGGACCCCGCTGCTTCTGCCA 326
Qy 241 ccggagcccaagcagccttggtcaggggtggtgtagcgtacaaaccttccagagtgtctgc 300
Db 327 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 386
Qy 301 taccaatatgtggacaccctatataccaggttttgaggggcacccgagatgtggaaccccaac 360
Db 387 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCCACCGAGATGTGGAACCCCAAC 446
Qy 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccggcctaca 420
Db 447 CGTGAGCTGACCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGACTACA 506
Qy 421 tccccaccctgtcctcgtcgtgatctatgggggtggttcttacagtggggcctcctcc 480
Db 507 TCCCCCAACCTGTCTCCTCGTCTGGATCATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 566
Qy 481 ttggacgtgtacgatggccgcttcttgggtacagggccgag 519

Db 567 TTGGACGTGTACGATGGCCCGCTTCTTGGTACAGGCCGAG 605
Search completed: August 31, 2002, 20:00:08
Job time: 16417 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:43:50 ; Search time 154.73 Seconds
(without alignments)
2738.433 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atgaggccccgcagtgtct.....aattgctcagcgccacctga 1725

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1722	99.8	1845	1 US-07-732-962A-1	Sequence 1, Appl
2	1722	99.8	1845	5 PCT-US92-06106-1	Sequence 1, Appl
3	1722	99.8	2256	2 US-08-318-826A-5	Sequence 5, Appl
4	1722	99.8	2256	2 US-08-370-156-1	Sequence 1, Appl
5	1722	99.8	2256	3 US-08-814-095-1	Sequence 1, Appl
6	1722	99.8	3016	2 US-08-318-826A-7	Sequence 7, Appl
7	1722	99.8	3016	2 US-08-370-156-5	Sequence 5, Appl
8	1722	99.8	3016	3 US-08-814-095-5	Sequence 5, Appl
9	1722	99.8	3096	2 US-08-318-826A-6	Sequence 6, Appl
10	1722	99.8	3096	2 US-08-370-156-3	Sequence 3, Appl
11	1722	99.8	3096	3 US-08-370-156-3	Sequence 3, Appl
12	1070	62.0	35060	3 US-08-814-095-7	Sequence 7, Appl
13	35	2.0	35	2 US-08-370-156-13	Sequence 13, Appl
14	35	2.0	35	2 US-08-370-156-23	Sequence 23, Appl
15	35	2.0	35	4 US-08-875-710-4	Sequence 4, Appl
16	20	1.2	20	3 US-08-990-065-4	Sequence 4, Appl
17	20	1.2	20	3 US-08-990-065-5	Sequence 5, Appl
18	20	1.2	20	3 US-08-990-065-7	Sequence 7, Appl
19	19	1.1	523	4 US-09-437-054A-5	Sequence 5, Appl
20	19	1.1	3159	4 US-09-437-054A-7	Sequence 7, Appl
21	19	1.1	246240	2 US-08-724-394A-20	Sequence 20, Appl
22	19	1.1	246240	2 US-08-724-394A-21	Sequence 21, Appl
23	19	1.1	246240	2 US-08-724-394A-22	Sequence 22, Appl
24	18	1.0	300	4 US-09-135-994-3	Sequence 3, Appl
25	18	1.0	477	2 US-09-109-266-17	Sequence 17, Appl
26	18	1.0	719	3 US-08-370-223-4	Sequence 4, Appl
27	18	1.0	940	1 US-08-353-400-24	Sequence 24, Appl

28	18	1.0	943	2 US-08-303-569B-4	Sequence 4, Appl
29	18	1.0	943	2 US-08-116-247-4	Sequence 4, Appl
30	18	1.0	1243	4 US-09-103-875-16	Sequence 16, Appl
c 31	18	1.0	1276	4 US-09-177-325-2	Sequence 2, Appl
c 32	18	1.0	1276	4 US-09-411-812A-2	Sequence 2, Appl
c 33	18	1.0	1276	4 US-09-590-113-2	Sequence 2, Appl
34	18	1.0	1701	3 US-09-264-737-3	Sequence 3, Appl
35	18	1.0	2184	1 US-08-445-050-8	Sequence 8, Appl
36	18	1.0	2184	1 US-08-204-691-8	Sequence 8, Appl
c 37	18	1.0	2301	1 US-08-306-691B-23	Sequence 23, Appl
c 38	18	1.0	2301	5 PCT-US93-06251-78	Sequence 78, Appl
39	18	1.0	2428	1 US-08-445-050-1	Sequence 1, Appl
40	18	1.0	2428	1 US-08-204-691-1	Sequence 1, Appl
41	18	1.0	2487	3 US-08-370-223-12	Sequence 12, Appl
42	18	1.0	3018	1 US-08-347-718B-3	Sequence 3, Appl
43	18	1.0	3018	1 US-08-482-262-3	Sequence 3, Appl
44	18	1.0	3018	6 5200183-1	Patent No. 5200183
c 45	18	1.0	3123	1 US-08-094-889-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-07-732-962A-1
; Sequence 1, Application US/07732962A
; Patent No. 5248604
; GENERAL INFORMATION:
; APPLICANT: Fischer, Melr
; TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
; TITLE OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,962A
; FILING DATE: 19910722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
US-07-732-962A-1

Query Match 99.8%; Score 1722; DB 1; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggcccccagtgctgctgcacacgccttccctggcttccccactccttctcctc 60
Db 1 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCCACTCTCTCCTC 60
QY 61 ctccctctgctccttggtggaagagtggtggggctgagggcccgaggatgcagagctgctg 120
Db 61 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 120
QY 121 gtgacggtgctggtggggcccgctgctggtgggcattegcctgaagacccccggggccctgtc 180
Db 121 GTGACGGTGCCTGGGGCCGGCTGCGGGGCAATCGCTGAAGACCCCCCGGGGCCCTGTC 180
QY 181 tctgcttccctgggcatccccctttgcggagccaccccatgggacccccgtcgtcttctgcca 240
Db 181 TCTGCTTCTCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGTCTGCTTCTGCCA 240
QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttccagagtctgc 300
Db 241 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTACAGCCTACAACCTTCCAGAGTGTCTGC 300
QY 301 taccaatatgtggacacccctataccccagggttttgagggccaccgagatgtggaaccccaac 360
Db 301 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 360
QY 361 cgtgagctgagcggagactgcctgtacctcaacgtgtggacaccataccccggccctaca 420
Db 361 CGTGAGCTGAGCGAGGACTGCCTGTACTACCTCAACGTGTGGACACCATAACCCCGGCCCTACA 420
QY 421 tccccacccctgtcctcgtctggtatctatgggggtggtcttctacagtgggcctcctcc 480
Db 421 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCCTCC 480
QY 481 ttggacgtgtacgtggccgcttcttctgttacaggccgagaggactgtgctggtgccatg 540
Db 481 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGTGTCTCCATG 540
QY 541 aactaccgggtggagccctttggcttccctggccctgccggggagccgagagcccgccggc 600
Db 541 AACTACCGGGTGGAGCCTTTGGCTTCTTGGCTTCCCTGGCCCTGCCGGGAGCCGAGAGCCCGGGC 600
QY 601 aatgtgggtctcctggatcagaggctggccctgcagtggtgagcagagaacgtggcagcc 660
Db 601 AATGTGGGTCTCTTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 660
QY 661 ttccgggggtgaccgcacatcagtgacgctgtttggggagagcgcgggagccgcctcggtg 720
Db 661 TTCGGGGGTGACCCGACATCAGTGACGCTGTTGGGGAGAGCGCGGGAGCCGCTCGGTG 720
QY 721 ggcattgcacctgctgtccccgccagccggtggttccacagggccgtgctgcagagc 780
Db 721 GGCATGCACCTGCTGTCCCCCGCCAGCCGGGGCTGTTCCACAGGGCCGCTGCTGCAGAGC 780
QY 781 ggtgcccccaatgacccctggccacggtgggcatgggagagcccgctgcagggccacg 840
Db 781 GGTGCCCCCAATGGACCCCTGGGCCACGCTGGGCATGGGAGAGGCCCCGTGCGAGGGCCACG 840
QY 841 cagctggccccacctgtgtgggtgtcctccaggcgggcactggtgggaatgacacagagctg 900
Db 841 CAGCTGGCCCCACCTGTGTGGGTGTCTCTCCAGSGCGCACTGGTGGGAATGACACAGAGCTG 900
QY 901 gtgacctgcttcggacacgaccagcagcaggtcctggtgaaccacgaaatggcacgtgctg 960
Db 901 GTAGCCTGCCCTTCGACACAGCACGACCGCAGGTCTGTTGAACCAACGAAATGGCAGCTGCTG 960
QY 961 cctcaagaaaagcgtcttcgggttctcctcgtgctggtggtagatggagaacttccctcagt 1020
Db 961 CCTCAAGAAAGCGTCTTCCGGTCTCTCTCTGCTGTGGTAGATGGAGACTTCTCTCAGT 1020
QY 1021 gacacccagaggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
Db 1021 GACACCCAGAGGCCCTTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGTGGGT 1080
QY 1081 gtggtgaaggataggggctcgtatttctgtttacggggcccccagggttcagcaagac 1140

Db 1081 GTGGTGAAGGATGAGGGCTCGTATTTCCTGGTTTACGGGGCCCCAGGCTTCAGAAAGAC 1140
QY 1141 aacgagctctcatcagccgggcccaggttccctggtgggggtgcgggtcggggttccccag 1200
Db 1141 AACGAGTCTCTCATCAGCCGGCCGAGTTCTTGCCCGGGGTGCGGGTCTCCCGAG 1200
QY 1201 gtaagtacctggcagcggaggtgtggtcctgcattacacagactggctgcatccccag 1260
Db 1201 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCGAG 1260
QY 1261 gacccggcacgcctgaggggagccctgagcgtggtggtggcgacacaaatgctcgtgtgc 1320
Db 1261 GACCCGGCACGCCCTGAGGGAGGCCCTGAGCGATGTGTTGGGCGACCAACAATGCTCGTGTGC 1320
QY 1321 cccgtggccccagtggtggcgactggctgccccgggtgccccgggtctacgcctacgtc 1380
Db 1321 CCCGTGGCCCCAGCTGGCTGGGCGACTGGCTGCCAGGGTGGCGGCTCTACGGCTACGTC 1380
QY 1381 ttgaaacacctgcttccacgctctcctggccccctgtggatgggggtgccccacgggtac 1440
Db 1381 TTTGAACACCGTGTTCACGCTCTCCTGGCCCCCTGTGGATGGGGTGGCCCCACGGCTAC 1440
QY 1441 gagatcgagttcatctttgggatccccctggacccctctcgaaactacacgagcagaggag 1500
Db 1441 GAGATCGAGTTCACTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGCGCAGAGGAG 1500
QY 1501 aaaaatttcgccccagcagactgatgcgatactgggccaactttgccccacagggggtccc 1560
Db 1501 AAAATCTTCGCCCAGCGACTGATCGGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1560
QY 1561 aatgagccccgagaccccaaggccccacaaatggcccccgctacacgcgggggtcagcag 1620
Db 1561 AATGAGCCCCGAGACCCCAAGGCCCCACAATGGCCCCCGTACACGGGGGGGCTCAGCAG 1620
QY 1621 tacgttagtctggacctgcggccgctggaggtggtggcggtgcgcgccagggcctgc 1680
Db 1621 TACGTTAGTCTGGACCTGCGGCGCGTGGAGGTGCGGGGGGGGTGCGCGCCAGGGCCTGC 1680
QY 1681 gccttctggaacccgcttccctcccccaaatgctcagcgccacc 1722
Db 1681 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGCCACC 1722

RESULT 2
PCT-US92-06106-1
; Sequence 1, Application PC/TUS9206106
; GENERAL INFORMATION:
; APPLICANT: Fischer, Meir
; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06106
; FILING DATE: 19920722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
PCT-US92-06106-1

Query Match 99.8%; Score 1722; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgaggccccgcagtgctgctgcacagccctccctgctccctccactcctcctc 60
Db 1 ATGAGGCCCCGCAGTGTCTGCTGCACAGCCCTCCCTGGCTTCCCACACTCTTCTCCTC 60

Qy 61 ctccctctgctcctgggtgagagtggtgggctgagggccgggaggtgcagagctgctg 120
Db 61 CTCCTCTGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 120

Qy 121 gtgacggtgctggtggggccggtgcgggggcatctgcctgaagacccccggggccctgtc 180
Db 121 GTGACGGTGGTGGGGCCGGCTCGGGGCATTCGCCCTGAAGACCCCCGGGGCCCTGTC 180

Qy 181 tctgcttctctgggcaccccccttggcagggccacccatgggaccccgctcgtctgtc 240
Db 181 TCTGCTTTCTCTGGGCATCCCCCTTTGCGAGGCCACCCATGGGACCCCGTCGCTTTCTGCCA 240

Qy 241 ccggagcccaagcagccttgctcaggggtggtagacgtacaaaccttcagagtgctctgc 300
Db 241 CCGGAGCCCAAGCAGCCTTGGTCAAGGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 300

Qy 301 taccaatatgtggacaccctatacceaggtttgagggcaccgagatgtggaaccccaac 360
Db 301 TACCAATAATGTGGACACCCTATACCCAGGTTTGAAGGACCCGAGATGTGGAACCCCAAC 360

Qy 361 cgtgagctgagcagagactgcctgtacctcaacgtgtggacaccataccccggcctaca 420
Db 361 CGTGAGCTGAGCGAGGACTGCCCTGTACTCAACGTGTGGACACCATACCCCCGGCCTACA 420

Qy 421 tccccacccccctgtcctcgtctggtatctatgggggtggtcttctacagtgggccctcctc 480
Db 421 TCCCCACCCCTGTCTCTGCTGTGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 480

Qy 481 ttggacgtgtacgatggccgcttcttggtacagggccgagagactgtggtgtccatg 540
Db 481 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 540

Qy 541 aactacgggtgggagccttggcttctcctggccctgcccggggagcagagagggcccgggc 600
Db 541 AACTACGGGTGGGAGCCCTTTGGCTTCTCGGCCCTGCCGGGAGCCGAGAGGCCCGGGC 600

Qy 601 aatgtgggtcctcctggatcagaggtgcccctgcagtgggtgcagagagaacgtggcagcc 660
Db 601 AATGTGGGTCTCCTGGATCAGAGGCTGSCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 660

Qy 661 ttcgggggtgaccccgacatcagtgacgtgtttggggagagcgcggggagccctcgtg 720
Db 661 TTCGGGGGTGACCCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCGCCCTCGGTG 720

Qy 721 ggcattgcacctgtctcccccgccagccggggcctgttccacagggcgtgctgcagagc 780
Db 721 GGCATGCACCTGTCTCCCGCCAGCCGGGGCCTGTTCCACAGGGCGGTGCTGCAGAGC 780

Qy 781 ggtgcccccaatggaccctgggccacgtgggcatgggagaggcccgctgcaggggccacg 840

Db 781 GGTGCCCCCAATGGACCCCTGGGGCCACGGTGGGCATGGGAGAGGCCCGTCCAGGGCCACG 840

Qy 841 cagctgccccacacttggtggtgtcctccagggcagctggtggaatgacacagagctg 900
Db 841 CAGCTGGCCCACTTGTGGGCTGTCTCCAGCGGCACCTGGTGGGAATGACACAGAGCTG 900

Qy 901 gtagcctgccttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
Db 901 GTAGCCTGCCTTCGGACACGACGACGCGAGCTCCTGGTGAACCCACGAATGCGACGTGCTG 960

Qy 961 cctcaagaaaagcgtcttcctcgttctcctcctcgtcctggtgtagatggagacttcctcagt 1020
Db 961 CCTCAAGAAAAGCGCTTTCGGGTTCTCCTTCTCCTTCTGCTGCTGTTAGATGGAGACTTCCTCAGT 1020

Qy 1021 gacacccccagaggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
Db 1021 GACACCCACAGAGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGTTGGT 1080

Qy 1081 gtggtgaaggtgagggctcgtattttctgttttacggggccccagggcttcagcaaaagac 1140
Db 1081 GTGTTGAAGGATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCAGGCTTCAGCAAAAGAC 1140

Qy 1141 aacgagtcctcatcagccggcgagttcctgcccgggtgcgggtgcgggttccccag 1200
Db 1141 AACGAGTCTCATCAGCCGGCCGAGTTCCTGGCCGGGTGCGGGTCCGGGTTCGCCAG 1200

Qy 1201 gtaagtgaacctggcagccgaggtgctggttcctgcattacacagactggctgcacatcccgag 1260
Db 1201 GTAAGTGACCTGCGACCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCCGAG 1260

Qy 1261 gacccggcacgcctgagggagggccctgagcgtggtggtggcgacacaaatgctgctgctgc 1320
Db 1261 GACCCGGCACGCCTGAGGGAGGCCCTGAGCGATGTGGTGGGGCCACCAATGTCTGTGTGC 1320

Qy 1321 cccgtgccccagctggctggcgactggctggcctggcctggcctggcctggcctggcctggcct 1380
Db 1321 CCGGTGGCCAGCTGGCTGGCGGACTGGCTGGCCAGGGTGCCCGGGTCTACGCCCTACGTC 1380

Qy 1381 tttgaacacccgtgcttccacgctcctcctggcctggcctggcctggcctggcctggcctggcct 1440
Db 1381 TTTGAACACCCGTGCTTCCACGCTCTCCTGGCCCTGTGGATGGGGTGCCCGGCTAC 1440

Qy 1441 gagatcgagttcatcttgggatccccctggacccccctctcgaaactacacgagagagag 1500
Db 1441 GAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGCGCAGAGGAG 1500

Qy 1501 aaaatcttcgccagcagctgatgcgatacttggtggccttgcctggcctggcctggcctggcct 1560
Db 1501 AAAATCTTCGCCCAGCGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1560

Qy 1561 aatgagccccgagagacccccaaaggcccccaaatggcccccgctacacggcgggggtcagcag 1620
Db 1561 AATGAGCCCCCGAGACCCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGGCTCAGCAG 1620

Qy 1621 tacgttagtctggaacctgcggccgctggaggtgcgggggtgcgggggtgcgggggtgcgggg 1680
Db 1621 TACGTTAGTCTGGACCTGCGCCGCTGGAGGTGCGGGCGGGGGTGCGCCGCCAGGGCCTGC 1680

Qy 1681 gccttctggaacccgcttccctcccccaaatgtcagcgccacc 1722
Db 1681 GCCTTCTGGAACCGCTTCTCCCCAAAATTGCTCAGCGCCACC 1722

RESULT 3
US-08-318-826A-5
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense

QY 1501 aaaatcttcgcccagcgactgatcgatacttgggccaaacttggcccgacacaggggatccc 1560
|||||
Db 1660 AAAATCTTGCCTCCAGCGACTGATGCGATACTGGGGCAACTTTTGCCCGCACAGGGGATCCC 1719

QY 1561 aatgagcccgagaccccaagccccacaatggcccccgctacacgaggggggctcagcag 1620
|||||
Db 1720 AATGAGCCCGAGACCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779

QY 1621 tacgttagctggacctgcccgcctggaggtgcggggctgcccagggcctgc 1680
|||||
Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGGGGGCTGCGGCCCCAGGCCTGC 1839

QY 1681 gccttctggaaccgcttcctcccaaatgctcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCTCCCCAAATGCTCAGCGCCACC 1881

RESULT 4

US-08-370-156-1
; Sequence 1, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethlington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:

; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-156-1

Query Match 99.8%; Score 1722; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggcccgagtgctgctgcacagccttccctggcttcccgacttctctc 60
|||||
Db 160 ATGAGGCCCCGCGAGTGTCTGCTGCACAGCGCTTCCCTGGCTTCCCACTCCTTCTCCTC 219

QY 61 ctcctctgctcctgggtgagaggtggggctgagggccgggagagatgcagagctg 120
|||||
Db 220 CTCCTCTGGTCTCTGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279

QY 121 gtgacgggtgcgtggggccggtgcggggcattgcctgaagacccccggggggccctgtc 180
|||||
Db 280 GTGACGGTGGTGGGGCCGGCTGCGGGGCAATTCGCTGAAGACCCCCGGGGGCCCTGTGTC 339

QY 181 tctgcttctcctgggcatcccccttgcggagcaccatgggacccccgtcttcttccca 240
|||||
Db 340 TCTGCTTCTCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGCTCTGTGCCA 399

QY 241 ccggagcccaagcagccttggtaggggtgtagacgctacaaccttccagagtgtctgc 300
|||||
Db 400 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGC 459

QY 301 taccaatatgtgacacccctataccagggttttgagggcaccgagatgtggaaccccaac 360
|||||
Db 460 TACCAATATGTGACACACCTATACCCAGGTATTGAGGGCACCGAGATGTGGAACCCCAAC 519

QY 361 cgtgagctgagcgagactgcctgtacctcaactggtggacaccatacccccgccctaca 420
|||||
Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCCCTACA 579

QY 421 tccccaccctcctcctcgtcgtggtatctatgggggtggtcttctacagtggggcctcctcc 480
|||||
Db 580 TCCCCCACCCCTGTCTCTGCTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTTCTCTCC 639

QY 481 ttggacgtgtacgatggcgccttcttggtacaggccgagaggactgtgctggtgtccatg 540
|||||
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGTGTCCATG 699

QY 541 aactaccgggtgggagccttggcttctcctggcctgcccctgcgggagccgagagggcccgggc 600
|||||
Db 700 AACTACCGGGTGGGAGCCTTTGGCTTCTTGGCCTTCCCTGGCCCTGCCGGGAGCCCGAGAGCCCGGGGC 759

QY 601 aatgtgggtcctcctggtatcagaggtgcctgcctgcagtggtgagaggaacgtggcagcc 660
|||||
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGSCCTGCAGTGGTGGTGCAGGAGAACTGGCAGCC 819

QY 661 ttcgggggtgacccgacatcagtgacgctgtttggggagagcgcgggagccgcctcgggtg 720
|||||
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTGGGGAGAGCGCGGGGAGCGCCCTCGGTG 879

QY 721 ggcacgacctgctgtccccgcgccagccgggctgttccacagggccgtgctgcagagc 780
|||||
Db 880 GGCATGCACCTGTGTCCCCGCCACAGCCGGGGCCTGTTCACAGGGCCGTGTGCAGAGC 939

QY 781 ggtgcccccaatggacccctgggccacggtgggcatgggagagccccgtcgcagggccacg 840
|||||
Db 940 GGTGCCCCCAATGGACCTTGGGCCACCGTGGCATGGGAGGCCCGCTCGAGGGGCCACG 999

QY 841 cagctggccacacttggtggctgctcctcagggcgccactggtgggaatgacacagagctg 900
|||||
Db 1000 CAGCTGGCCCACTGTGGGTGTCTCCTCCAGCGGCACTGGTGGGAATGACACAGAGCTG 1059

QY 901 gtagcctgccttcgggacacgacacgagcagctcctggtgaaccacgaatggcacgtgctg 960
|||||
Db 1060 GTAGCCTGCCCTCGGACACGACCGCAGCGCAGTCTCTGTTGAACCCAGTAATGGCAGCTG 1119

QY 961 cctcaagaaaaagcgtcttccgggttctcctcctgctgctggtgtagatggagacttccctcagt 1020
|||||
Db 1120 CCTCAAGAAAGCGCTTTCGGGTCTCTCTTCTGCTGTGGTAGATGGAGACTTCTCTCAGT 1179

QY 1021 gacaccccgagggccctcatcaacgcggggagacttccacggcctgcagtgctggtgggt 1080
|||||
Db 1180 GACACCCCGAGAGGCCCTCATCAACGCGGGGAGACTTCCACGGCCTGCAGGTGTGGTGGGT 1239

QY 1081 gtggtgaaggatgagggctcgtattttctggttttacggggccccagggttccagcaaaagac 1140
|||||
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCAGGCTTCCAGCAAGAC 1299

QY 1141 aacgagtcctctcatcagccggggccgagttccttggccggggtgcgggtgtcccccag 1200
|||||
Db 1300 AACGAGTCTCTCATCAGCCGGGCCGAGTTCTTGGCCGGGTGCGGGGTGCGGGGTTCCTCCAG 1359

QY 1201 gtaagtgcacctgggcagccgaggtggtggtcctgcattacacagactggtgcatccccag 1260

|||||
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCGAG 1419
QY 1261 gacccggcacgcctgagggagggccctgagcgaatggtggtggcgaccacaaatgctgtgtgc 1320
|||||
Db 1420 GACCCGGCACGCTGAGGGAGGCCCTGAGCGGATGTGGTGGCGGACCAACAATGTCTGTGTGC 1479
QY 1321 cccgtgccccagctggctggcgactggctgccccaggggtgccccgggtctacgcctacgtc 1380
|||||
Db 1480 CCGGTGGCCACGCTGGCTGGCGGACTGGCTGCCAGGGTGCCCGGGTCTACGCTACGTC 1539
QY 1381 ttgaaacaccgtgtctccacgctctcctggccccctgtggtgggtgccccacggctac 1440
|||||
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCCTGGCCCTGTGGATGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcgagttcatctttgggatccccctggacccctctcgaaaactacacggcgagagag 1500
|||||
Db 1600 GAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaatcttgcgccagagactgatgcgatactggtggccaaactttgccccacaggggatccc 1560
|||||
Db 1660 AAAATCTTCGCCCAGCAGCTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagaccacaagggccccacaatggccccctgtacacggcgggggtcagcag 1620
|||||
Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtctggacctgcccgcgctggaggtgcggcgggggtgcgcgcgccagggcctgc 1680
|||||
Db 1780 TAGCTTAGTCTGGACCTGCGGCCGCTGGAGGTGGCGGGGGGCTGCGCGCCAGGCCCTGC 1839
QY 1681 gccttctggaaccgcttctctccccaaaattgctcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCCTCCCCAAATTGCTCAGGCCACC 1881

RESULT 5

US-08-814-095-1
; Sequence 1, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2256 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "ACHE gene comprising exons
; DESCRIPTION: 2, 3, 4 and 6"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-814-095-1

Query Match 99.8%; Score 1722; DB 3; Length 2256;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaggccccgcagtgctctgctgcacacgccttccctggcttccccactccttctcctc 60
|||||
Db 160 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCCTTCCTGGCTTCCCCACTCCTTCTCCTC 219
QY 61 ctctctggtctcctgggtggaggagtgggggtgagccctgcctgaagacccccggggccctgtc 120
|||||
Db 220 CTCTCTGGCTCCTGGTGGAGGAGTGGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 279
QY 121 gtacggtgcgtggggccggctgcgggggcattgcctgaagacccccggggccctgtc 180
|||||
Db 280 GTGACGGTGCCTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCCGGGGCCCTGTC 339
QY 181 tctgcttctcctgggcatccccctttgcggagccaccatggacccctgccttcttctgcc 240
|||||
Db 340 TCTGCTTTCTTGGGCATCCCCCTTTGCGGAGCCACCATGGGACCCCCGCTGCTTCTGCCA 399
QY 241 ccggagcccaaagccttggtcaggggtggtagacgcctacaaccttccagagtgtctgc 300
|||||
Db 400 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACCTACAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacacccctatacccgaggttttgggggcacccgagatgtggaaccccaac 360
|||||
Db 460 TACCAATATGTGGACACCCCTATATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcgagactgcctgtacctcaacgtgtgtgacacccatacccccgccctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGCTGTGGACACCATACCCCGGCTACA 579
QY 421 tccccacccctgtcctgctggtgatctatgggggtggttcttacagtggggctcctcc 480
Db 580 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCTCC 639
QY 481 ttggacgtgtacgatggcgcttcttgggtacaggccgagagagactgtgctggtgtccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGSCCGGAGGACTGTGCTGGTGTCCATG 699
QY 541 aactaccgggtgggagccttggcttccctggcctgcgggggagccgagagagccccgggc 600
Db 700 AACTACCGGTGGGAGCCTTTGGCTTCTTGGCCCTTGGCGGGAGCCGAGAGGCCCGGGGC 759
QY 601 aatgtgggtctcctggaatcagaggtggtgccccctgagtggtgaggaacgtggcagcc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttcgggggtgacccgacatcagtgacgctgtttggggagagcgcgggagccgctcgggtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGGAGCCGCTCGGTG 879
QY 721 ggcattgcacctgctgtccccgccccagccggggcctgttccacagggccgtgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCCGCCACGCCGGGGCCCTGTCCACAGGGCCGCTGCTGCAGAGC 939
QY 781 ggtgcccccaatggacctggccacggtgggcatgggagagggccctgcgagggccacg 840
Db 940 GGTGCCCCCAATGGACCTTGGGCCACCGTGGGGCATGGGAGAGGGCCCGCTCGCAGGGCCACG 999

QY 841 cagctggcccaacctgtgtgggtgtctctccagggcagggcactggtgggaatgacacagagctg 900
|||||
Db 1000 CAGCTGGCCCCACCTTGTGGGCTGTCCTCCAGGGCGCACTGGTGGGAATGACACAGAGCTG 1059

QY 901 gtagctgccttcggacacgacgacgcagctcctggtgaaccacgaatggcacgtgctg 960
|||||
Db 1060 GTAGCCTGCCTTCGGACACGACGACGCGCAGGTCCTGGTGAACCACGAATGCACGTGCTG 1119

QY 961 cctcaagaaagcgtcttccttccttccttcctgcctggtgtagatggagacttccctcagt 1020
|||||
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCTCCTTCGTGCCCTGTGGTAGATGGAGACTTCCTCAGT 1179

QY 1021 gacacccccagagggccctcatcaacgcgggagagacttccacggcctgcaggtgctggtgg 1080
|||||
Db 1180 GACACCCACAGAGGCCCTCATCAACGCGGGGAGACTTCCACGGGCTGCAGGTGCTGGTGGGT 1239

QY 1081 gtggtgaaggatgagggctcgtaatttctgtgtttacggggccccaggcttcagcaaaagac 1140
|||||
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTCTGTGTTTACGGGGCCCCAGGCTTCAGCAAAGAC 1299

QY 1141 aacgagtctctcatcagccgggcccagagttcctgcccggggtgcgggttccccag 1200
|||||
Db 1300 AACGAGTCTCTCATCAGCCGGCGGAGTTCCTGGCCGGGTGCGGGTTCGCCAG 1359

QY 1201 gtaagtgcctggcagccgagggcctgagcgtggtcctgcattacacagactggctgcacccgag 1260
|||||
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCGATTACACAGACTGGCTGCATCCCGAG 1419

QY 1261 gaccgggcacgcctgagggagggcctgagcgtggtggtggcgaccacaaatgctgctgc 1320
|||||
Db 1420 GACCCGGCACGCCCTGAGGGAGGCCCTGAGCGATGTGGTGGGGGACCCACAATGTCGTGTGC 1479

QY 1321 cccgtggccccagctggctggggcgaactggctgcccagggtgcccgggtctacgcctacgtc 1380
|||||
Db 1480 CCCGTGGCCCCAGCTGGCTGGGGGACTGGCTGCCAGGGTGCCCGGGTCTACGCCCTACGTC 1539

QY 1381 ttgtgaacacctgtctccacgctctctctgcccctgtggatgggggtgccccacggctac 1440
|||||
Db 1540 TTTGAACACCGTGTCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599

QY 1441 gagatcagattcatctttgtgggtacccccctggacccccctctcgaactacacgagaggag 1500
|||||
Db 1600 GAGATCAGATTCTATTGGGATCCCCCTGGACCCCCTCTCGAAACTACACGGCAGAGGAG 1659

QY 1501 aaaaattctgccccagcagcagcccccaaggcccccaaatggcccccgtaacagggcgctcagcag 1560
|||||
Db 1660 AAAATCTTGGCCCCAGGACTGATGCGATACTTGGGCCAACATTGGCCCCGACAGGGGATCCC 1719

QY 1561 aatgagcccccgagaccccccaaggcccccaaatggcccccgtaacagggcgctcagcagcag 1620
|||||
Db 1720 AATGAGCCCCGAGACCCCCAAGGCCCCCAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779

QY 1621 tacgttagtctggacctgcggccgctggaggtgcggcggggctgcgcgccccagggcctgc 1680
|||||
Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGGCGGGGGCTGCGCGCCCGAGGCGCTGC 1839

QY 1681 gccttctggaaccgcttctctcccaaatgtctcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGGCCACC 1881

RESULT 6
US-08-318-826A-7
; Sequence 7, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; OTHER INFORMATION: /note= "Splice Variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4, 5 and 6"
US-08-318-826A-7

Query Match 99.8%; Score 1722; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctgtgtgcacacgcttccctggcttccccactctctctc 60
|||||
Db 160 ATGAGGCCCCCGCAGTGTCTGTGCACACGCCCTTCCCTGGCTTCCCACACTCTCTCCTC 219

QY 61 ctctctggtcctgggtggtgaggtggtggggctgagggccgggaggtgagagctgctg 120
|||||
Db 220 CTCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGACAGAGTGTCTG 279

QY 121 gtgacggtgcgtggggccggctgcggggcattcgctgaagacccccggggccctgtc 180
|||||
Db 280 GTGACGGTGCCTGGGGCCCGGCTGCGGGGCATTCGCTGAAGACCCCGGGGCCCTGTC 339

QY 181 tctgcttctctgggcatccccctttgcggagaccacccatgggaccccgctcttctgcca 240
|||||
Db 340 TCTGCTTTCCTGGGCATCCCCCTTTGCGAGGCCACCCATGGGACCCCGTCGCTTCTGCCA 399

QY 241 ccggagcccaagcagccttggtaagggtggttagacgctacaacctccagagtgtctgc 300
|||||
Db 400 CCGAGGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 459

QY 301 taccaatatgtggacacccctataccccaggttttggaggcaccgagatggaaccccaac 360
|||||
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGGGGCACCCGAGATGTGGAAACCCCAAC 519

QY 361 cgtgagctgagcaggtgactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
|||||
Db 520 CGTGAGCTGACCGAGGACTGCCCTGTACCTCAACGTTGTGGACACCCATACCCCGGCCCTACA 579

QY 421 tccccacccctgtctcgtctggtatctatgggggtggtcttctacagtggggcctcctcc 480
Db 580 TCCCCACCCCTGTCCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCCTCC 639
QY 481 ttggacgtgtacatggccgcttcttcttggtacagccgagaggactgtctggtgtccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGTGGTGTCCATG 699
QY 541 aactaccgggtggagaccttggcttctcctggccctgccggggagccgagagggcccgggc 600
Db 700 AACTACGGGTGGAGCCCTTTGGCTTCCCTGGCCCTGCCGGGAGCCGAGAGCCCCCGGGC 759
QY 601 aatgtgggtctctggatcagagggctggccctgcagtggtgaggtgcaggagaacgtggcagcc 660
Db 760 AATGTGGGTCTCTTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttcgggggtgaccgacatcagtgcgctgtttggggagagcgcgggacctcggtg 720
Db 820 TTCGGGGTGACCCGACATCAGTGACGCTGTGGGAGAGCGCGGGAGCGCCTCGGTG 879
QY 721 ggcatacctgtgtccccgccagccgggacctgttccacagggccgtgtgcagagc 780
Db 880 GGCATGCACCTGTCTCCCGCCAGCCGGGCTGTTCACAGGGCCGTGTGCAGAGC 939
QY 781 ggtgcccccaatgaccctgggcccacggtgggcatggagagggcccgagggccacg 840
Db 940 GGTGCCCCCAATGACCCCTGGGCCACGGTGGCATGGAGAGCCCGTCCAGGGCCACG 999
QY 841 cagctggccccacctgtgggctgtcctccagggcgacctggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCCCACTGTGGGCTGTCTCTCCAGCGGCACACTGGTGGGAATGACAGAGCTG 1059
QY 901 gtagecctgccttgggacacgaccagcgcaggtcctggtgaaccacgaaatggcacgtgctg 960
Db 1060 GTAGCCTGCCCTTCGGACACGACGACGACGCGGCTGCTGGTGAACCAAGATGGCAGCTGTG 1119
QY 961 cctcaagaaagcgtcttccgggttctcctcctcgtgctgtgtagatggagacttccctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTCCGGTTCCTCTCGTGCCTGTGGTAGATGGAGACTTCCTCAGT 1179
QY 1021 gacacccagagccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
Db 1180 GACACCCAGAGGCCCTCATCAACGCGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT 1239
QY 1081 gtggtgaaggatagggctcgtatttctggttacggggcccccaggcttcagcaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTCTGCTTTACGGGGCCCCAGGCTTCAGCAAGAC 1299
QY 1141 aacgagtctctcatcagccgggcccaggttctcctggcggggtgcgggtggttccccag 1200
Db 1300 AACGAGTCTCTCATAGCCGGGCGGAGTTCTTGGCCGGGTGCGGGTTCGGGTTCCCCAG 1359
QY 1201 gtaagtgaacctggcagccgaggtgtggtcctgcattacacagactggctgcacatcccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGGTGCATCCCCGAG 1419
QY 1261 gacccggcacgcctgagggaggccctgagcgtgtggtggcgaccacaatgtcgtgtgc 1320
Db 1420 GACCCGGCACGCCTGAGGGAGGCCCTGAGCGATGTGGTGGCGGACCAACAATGTCTGTGTC 1479
QY 1321 cccgtggccagctgggtggcgactggctggccaggggtgccgggtgtacgcctacgtc 1380
Db 1480 CCGTGGCCAGCTGGCTGGCGACTGGCTGGCCAGGGTGCCTGGGTTCACGCCCTACGTC 1539
QY 1381 ttgaaacaccgtgtctccacgctcctcctggccctgtggatgggggtgccccacggctac 1440
Db 1540 TTTGAACACCCGTGTCTCCACGCTCTCTCTGGCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcgagttcatctttgggatccccctggacccctctcgaaaactacacgggcagaggag 1500
Db 1600 GAGATCGAGTTCACTTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659

QY 1501 aaaatcttcgcccagcgactgatcgatactgggccaactttgccccgcacaggggatccc 1560
Db 1660 AAATCTTCGCCACGCGACTGATCGGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagaccccccaaggccccacaatagccccctgtacacggcgggggtcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtctgacctgcggccgctggaggtgcgggggctgcgcgccagggcctgc 1680
Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGGGGCGGGGCTGCGCGCCAGGCCTGC 1839
QY 1681 gccttctggaaccgcttctcctccccaaaattgctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCCTCCCCAAATTGCTCAGCGCCACC 1881

RESULT 7
US-08-370-156-5
; Sequence 5, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
US-08-370-156-5

Query Match 99.8%; Score 1722; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagggccccgcagtgctgtgctgcacacgccttccctggcttccccactcttctcctc 60
Db 160 ATGAGGGCCCCCGCAGTGTCTGCTGCACACGCGCTTCCTGGCTTCCCCACTCTCTCCTC 219
QY 61 ctcctctggtcctcctggtgaggtggtgggggtgagggccgggaggtgcagagagctgctg 120

Db 220 CTCCTCTGGCTCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279

Qy 121 gtacggtgcgtggggccggctgggggcatcgcctgaagacccccggggccctgtc 180

Db 280 GTGACGCTGCTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGCCCTGTC 339

Qy 181 tctgcttctcctgggcatccccctttgaggagccaccatgggaccccgctgcgttctgtcca 240

Db 340 TCTGCTTCTCTGGCATCCCCCTTTGGGGAGCCACCCATGGGACCCCGTCGCTTCTGCCA 399

Qy 241 ccggagcccaagcagccttggtcagggggtggtagacgctacaaccttccagagtgtctgc 300

Db 400 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 459

Qy 301 taccaatatgtggacaccctataccacaggttttgagggcaccgagatgtggaaccccaac 360

Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGGATGTGGAACCCCAAC 519

Qy 361 cgtgagctgagcgaggactgcctgctggatcctgaacctcaactgtgtggacacataccccggcctaca 420

Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTTGTGGACACCATACCCCGGCCTACA 579

Qy 421 tccccacacccctgtcctcgtcctggatcctatgggggtggccttctacagtggggcctcctcc 480

Db 580 TCCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 639

Qy 481 ttggacgtgtacgatggccgctctcttggttacagggccgagagagactgtcgtgtccatg 540

Db 640 TTGGACGTTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGTGTCTCCATG 699

Qy 541 aactaccgggtgggagcctttggcttctcctggcctgcctggccctgcgggagcagagggccccgggc 600

Db 700 AACTACCGGTTGGGAGCCTTTGGCTTCTCGGCCCTGCCGGGAGCGGAGAGGCCCCCGGC 759

Qy 601 aatgtgggtcctcctggatcagaggtgcctcctgcctgcagtggtgcagagaaactggcagacc 660

Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819

Qy 661 ttcgggggtgaccccgacatcagtgacgtgttggggagagcgcgggagccgctcggtg 720

Db 820 TTCGGGGGTGACCCGACATCAGTACGCTGTTTGGGAGAGCGCGGAGCCGCTCGGTG 879

Qy 721 ggcattgcacctgtgtcccccgccagccggggcctgttccacagggccgctgctgcagagc 780

Db 880 GGCATGCACCTGCTGTCCCCCGCCAGCCGGGGCCTGTTCCACAGGGCCGCTGTCAGAGC 939

Qy 781 ggtgcccccaatggaccctgggccacggtgggcatgggagagggcccgctgcaggggccag 840

Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCGTCGCAGGCCACG 999

Qy 841 cagctggcccaacctgtgggctgtcctcagggcgacactggtgggaatgacacagagctg 900

Db 1000 CAGCTGGCCACACTGTGGGCTGTCTCTCAGGGCGCACTGGTGGGAATGACACAGAGCTG 1059

Qy 901 gtagcctgccttcggacacgaccagcagcagggtcctggtgaaccacgaatggcacgtgctg 960

Db 1060 GTAGCCTGCCTTCGGACACGACACGCGCAGGTCCTGGTGAACCCACGAATGGCAGTGTG 1119

Qy 961 cctcaagaaagcgtcttccggttctcctcctcgtgcctggtggtagatggagacttccctcagt 1020

Db 1120 CCTCAAGAAACGCTCTCCGGTTCTCTCTCGTGCCTGTGGTAGATGGAGACTTCTCTCAGT 1179

Qy 1021 gacacccccagggccctcatcaacgcgggagacacttccacggcctgcaggtgctggtgggt 1080

Db 1180 GACACCCACAGAGGCCCTCATCAACGGGGGAGACTTCCACGGCCTGCAGGTGCTGGGT 1239

Qy 1081 gtggtgaaggatgagggtcgtctatttcttggtttacggggccccaggttccagcaagac 1140

Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1299

Qy 1141 aacgagctctcatcagccggggccgagttcctggccggggtgcgggttccccag 1200

Db 1300 AACGAGTCTCTCATAGCCGGCCGAGTTCCTGGCCGGGTGCGGGTTCGGGTTCGCCAG 1359

Qy 1201 gtaagtgaacctggcagccgagggctgtggtcctgcattacacagagctggctgcaccccgag 1260

Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGCTCCTGCTTACACAGACTGGCTGCATCCCGAG 1419

Qy 1261 gaccggcacgcctgagggagggccctgagcgtatggtggggcaccacaatgtcgtgtgc 1320

Db 1420 GACCCGGCAGCCTGAGGGAGGCCCTGAGCGATGTGTTGGCGACACACAATGTCTGTGC 1479

Qy 1321 cccgtggccagctggctggcgactcctcctggcctgcccaggggtcccgggtctacgcctacgtc 1380

Db 1480 CCCGTGGCCAGCTGGTGGCGACTGGCTGCCAGGGTCCCCGGGTCTACGCTACGCTACGTC 1539

Qy 1381 ttggaacacccgtgttccacgctcctcctggcctgccccctgtggatgggggtgccccacggctac 1440

Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTCTGGCCCCCTGTGGATGGGGTGGCCCCACGGCTAC 1599

Qy 1441 gagatcgagttcattctttgggatccccctggacccctctcgaaaactacacggcagaggag 1500

Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659

Qy 1501 aaaattctgcgccagcactgatgcatactggcccaactggccccctgcacaggggggtcagcag 1560

Db 1660 AAAATCTTCGCCCAGCGACTGATGCGATACTGGGCCAACTTTGCCGCACAGGGGATCCC 1719

Qy 1561 aatgagcccgagagaccccaaggcccccaaatggccccctgcacaggggggtcagcag 1620

Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCCAACAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779

Qy 1621 tacgttagtctggacctgcggcgcctgaggtgcggcgggggtgcgcgccccagggcctgc 1680

Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGGGGGGTGCGGCCCCAGGCCCTGC 1839

Qy 1681 gccttctggaaccgcttctcctccccaaaattgctcagcgcacc 1722

Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGCCACC 1881

RESULT 8

US-08-814-095-5

; Sequence 5, Application US/08814095

; Patent No. 6025183

; GENERAL INFORMATION:

; APPLICANT: Soreq, Hermona

; APPLICANT: Zakut, Haim

; APPLICANT: Shani, Moshe

; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR

; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: KOHN & ASSOCIATES

; STREET: 30500 No. 6025183thwestern Highway, Suite 410

; CITY: Farmington Hills

; STATE: Michigan

; COUNTRY: U.S.

; ZIP: 48334

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/814,095

; FILING DATE:

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Montgomery, Ilene N.

; REGISTRATION NUMBER: 38,972

; REFERENCE/DOCKET NUMBER: 2391.00066

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (248) 539-5050

; TELEFAX: (248) 539-5055

; INFORMATION FOR SEQ ID NO: 5:


```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced AChE
; ORIGINAL SOURCE: comprising exons 2, 3, 4, 5 and 6"
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
US-08-814-095-5

Query Match          99.8%; Score 1722; DB 3; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagcccccgcagtgctgtgcacacgccttccttgcttccctgcttcccaactcttctctc 60
Db 160 ATGAGGCCCCCGCAGTGCTGCTGTCACACGCCCTTCCCTGGCTTCCCCACTCCTTCTCCTC 219

QY 61 ctccctctggtcctcctgggtgagagtggtgggctgagggccgggaggtgagatgcagagctgtg 120
Db 220 CTCCTCTGCTTCCTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279

QY 121 gtgacggtgcgtggggccggtgcggggcattcgccctgaagacccccggggccctgtc 180
Db 280 GTGACGGTGCgtggggccggtgcggggcattcgccctgaagacccccggggccctgtc 339

QY 181 tctgctttcttgggcatccccctttgcggagccaccccatgggacccccgtcgtcttctgcca 240
Db 340 TCTGCTTTCTTGGGCATCCCCCTTTGCGGAGCCACCCCATGGGACCCCGCTCGCTTCTGCCA 399

QY 241 ccggagccccaagcagccttggtcaggggtggtgtagacgctacaaccttccagagtgctgc 300
Db 400 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGCTGC 459

QY 301 taccaatatgtggacacccctataccccaggttttgagggcaccgagatgtggaaccccaac 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAAACCCCAAC 519

QY 361 cgtgagctgagcagagactgcctgtacctcaacgctgtggacaccataccccccggcctaca 420
Db 520 CGTGAGCTGACGGAGGACTGCCTGTACTCTCAACGTGTGGACACCATACCCCCGGCCTACA 579

QY 421 tccccacacctgtcctcgtctggaatcatatgggggtggttcttctacagtggggcctcctcc 480
Db 580 TCCCCACACCCCTGTCTCTGCTCGTATCTATGGGGGTGGCTTCTACAGTGGGGCCTCTCTCC 639

QY 481 ttggacgtgtacgatggccgcttcttggtagacggccgagagactgtgctggtgtccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTAGACGCCGAGAGGACTGTGCTGGTGTCCATG 699

QY 541 aactaccgggtgggagccttttgcttccctgcccctgcccgggagccgagagggcccgggc 600
Db 700 AACTACCGGGTGGGAGCCTTTTGGCTTCTTGGCCCTGCCCGGGAGCCGAGAGGCCCCCGGGC 759

QY 601 aatgtgggtcctcctggatcagagaggtggccctgcagtggtgggtgcaggagaaactgtggcagcc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819

QY 661 ttccgggggtgaccccgacatcagtgacgctgttttggggagagcgcgggagccgcctcggtg 720
Db 820 TTCGGGGGTGACCCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGGAGCCGCTCGGTG 879

QY 721 ggcattgcacctgctgtccccgccagccgggggcctgttccacagggccgtgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCCGCCAGCCGGGGCCTGTTCCACAGSGCCGTGCTGCAGAGC 939

QY 781 ggtgcccccaaatggacccttggggccacggtgggcatgggagagggcccgctgcagggccacg 840
```

```
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGAGAGGCCCGTCGCAGGGCCACG 999
QY 841 cagctggcccacacttgggtgtcctcctcagggcggcactgggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCCCACCTTGTGGCTGTCTCTCCAGCGGCACACTGGTGGGAATGACACAGAGCTG 1059
QY 901 gtagcctgccttcggacacgaccagcgcaggctcctggtgaaccacgaatggcacgtgctg 960
Db 1060 GTAGCCTGCCTTCGGACACGACCAGCGCAGGTCTCTGTTGAACCCACGAATGGCACGTGCTG 1119
QY 961 cctcaagaagaagcgtcttcccggttctcctcgtgcctgtggtagatggagacttccctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCTCTCTCGTGCCTGTGGTAGATGGAGACTTCTCTCAGT 1179
QY 1021 gacacccagagggccctcatcaacgcgggagacttccacggcctcaggtgctggtgggt 1080
Db 1180 GACACCCCGAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGTTGGGT 1239
QY 1081 gtggtgaaggatgagggctcgtattttcttggttttacggggccccaggttcagcaaaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCAGGCTTCAGCAAAAGAC 1299
QY 1141 aacgagctctctcatcagccgggcccaggttctcctgcccgggtgcgggttcgggttccccag 1200
Db 1300 AACGAGTCTCTCATCAGCCGGGCCGAGTCTCTGGCCGGGTGCGGGTTCGGGTTTCCCCAG 1359
QY 1201 gtaagtgccttgcagccgaggtggtgctcctgcattacacagactggctgcacatcccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGATTTACACAGACTGGCTGCATCCCGAG 1419
QY 1261 gacccggacgcctgagggagggccctgagcgtgtggtggcgacacaaatgctgtgtgc 1320
Db 1420 GACCCGGCACGCTGAGGGAGGCCCTGAGCATGTGGTGGCGACCAACAATGCTGTGTGC 1479
QY 1321 cccgtggcccagctggtggcgactggtgcccaggtgcccgggtctacgcctacgtc 1380
Db 1480 CCCGTGGCCAGCTGGCTGGCGACTGGTGTCACAGGTGCCGGGTCTACGCCCTACGCTC 1539
QY 1381 ttgaaacacgctgtctccacgctctcctctctggccccctgtggatgggggtgccccacggctac 1440
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTTGCCCTCTCTTGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcgagttcatctttgggatccccctctggacccctctcgaaactacacacgagagag 1500
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaatcttcgccccagcactgatgcatacttgggccaactttgcccgcacaggggataccc 1560
Db 1660 AAAATCTTCGCCCCAGCAGCTGATGCGGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagacccccaaaggcccccaaatggccccctacacggggggtcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCCAAAGGCCCCCAATGGCCCCCGCTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtctggacctgcccgcgtgaggtgcgggggtgcgcgggggtcagccagcctgc 1680
Db 1780 TACGTTAGTCTGGACCTCGGCCGCTGGAGGTGCGGGGGGCTGCGCGGGGCTGCGCGCCCTGC 1839
QY 1681 gccttctggaacccgcttctctcccccaaatgtcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCCAAATTTGCTCAGCGCCACC 1881
```

```
RESULT 9
US-08-318-826A-6
; Sequence 6, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
```


; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4, 5 and the translated portion of Intron 4 (readthrough)"
US-08-318-826A-6

Query Match 99.8%; Score 1722; DB 2; Length 3096;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggcccccagtgctgctgcacacgccttcctggcttccccactccttcctc 60
Db 160 ATGAGGGCCCCGAGTGTCTGCTGCACACGCTTCCCTGGCTTCCCCACTCCTTCTCCTC 219

QY 61 ctccctctggctcctgggtggaggagtggggctgagggccggaggatgcagagctgtc 120
Db 220 CTCCTCTGGCTCCTTGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279

QY 121 gtgacggtgcgtggggccggctgcggggcattcgcctgaagacccccggggccctgtc 180
Db 280 GTGACGGTGC GTGGGGCCGGCTGCGGGGCATTTCGCCCTGAAGACCCCCGGGGCCCTGTC 339

QY 181 tctgctttcctgggcataccccttgcggagccaccatgggaccccgctcgtcttctgcc 240
Db 340 TCTGCTTTCCTGGGCATCCCTTTGCGGAGCCACCCATGGGACCCCTCGCTTTCCTGCCA 399

QY 241 cggagcccaagcagccttggtcaggggtggtagacgtacacacccctccagagtgtctgc 300
Db 400 CCGAGCCCCAAGCAGCCCTTGCTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 459

QY 301 taccaatatgtggacacccctatacccagggttttgagggcaccgagatgtggaaccccaac 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTTTGAGGGSCACCGAGATGTGGAACCCCAAC 519

QY 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccggccctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCCTACA 579

QY 421 tccccacccctgtcctcgtctggatctatgggggtggccttctacagtggggcctcctcc 480
Db 580 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 639

QY 481 ttggacgtgtacgatggccgcttcttgggtacagggccgagaggaactgtgtgtgtccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGTGTCTCCATG 699

QY 541 aactacgggtggagccttbgcttccctggcctgcccctgcccggagccgagagccccgggc 600
Db 700 AACTACCGGGTGGAGCCCTTTGGCTTCTTGCCCTGCCCGGAGCCGAGAGGCCCGGGC 759

QY 601 aatgtgggtcctcctggatcagaggtggccctgcagtggtgaggaagaactgtggcagcc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGAGGAGAACGTGGCAGCC 819

QY 661 ttcggggtgacccgacatcagtgcagcgtgtttggggagagcgcgagccctcggtg 720
Db 820 TTTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCCGCTCGGTG 879

QY 721 ggcattgcacctgtgtcccccagccagccggggcctgttccacagggccgtgtgcagagc 780
Db 880 GGCATGCACCTGTGTCCCCGCCACGCCGGGCCCTGTTCCACAGGGCCCGTGTGCAGAGC 939

QY 781 ggtgcccccaatggaccctgggcaacggtggcattgggagggcccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCGCTCGCAGGGCCACG 999

QY 841 cagctggcccaccttgtgggtctcctccagggcactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCACACCTTGTGGGTCTCCTCCAGCGGCACCTGTTGGGAATGACACAGAGCTG 1059

QY 901 gtagcctgccttcggacacgaccagcgaggtccctggtgaaccacgaatggcacgtgctg 960
Db 1060 GTAGCCTGCCTTCGGACACGACCGACGCGCAGGTCTCTGGTGAACACGAATGGCACGTGCTG 1119

QY 961 cctcaagaagaagcgtcttccgggttctccttccctgcctgtgtgtagatggagacttccctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTCTCCTTCTGTCCTGTGTGTAGATGGAGACTTCTCCTCAGT 1179

QY 1021 gacaccccgagaggccctcacaacgcgggagagacttccacgcctgcaggtgctggtgggt 1080
Db 1180 GACACCCCGAGAGGCCCTCATCAACCGGGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT 1239

QY 1081 gtcgtgaaggatgagggctcgtatttctggtttacggggcccccaggcttcagcaaaagac 1140
Db 1240 GTGTTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCCAGGCTTTCAGCAAAAGAC 1299

QY 1141 aagagctctctcatcagccgggcccagagttcctgcccgggtgcgggtgcgggtccccag 1200
Db 1300 AACGAGTCTCTCATCAGCCGGGCCGAGTTCTTGCCCGGGGTGCGGGGTTCCTCCAG 1359

QY 1201 gtaagtgaacctggcagccgaggtgtggtcctgcattacacagactggctgcataccccag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCCGAG 1419

QY 1261 gacccggcacgcctgagggagggccctgagcgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgc 1320
Db 1420 GACCCGGCACGCCTGAGGGAGGGCCCTGAGCGGATGTGGTGGGGGACCAACAATGTGCGTGTGC 1479

QY 1321 ccggtggcccgagctggctggcgactggctgcccagggtgcccgggttctacgcctacgtc 1380
Db 1480 CCGTGGCCCCAGCTGGCTGGCGACTGGCTGCCAGGGTGCCTGGGGGTCTACGCCCTACGTC 1539

QY 1381 ttgaaacacccgtgtctccacgtctcctggtggccctgtggtatgggggtgccccacgggtac 1440
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCCTGGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599

QY 1441 gagatcgagttcatcttttgggatccccctggacccccctctcgaaactacacggcagaggag 1500

```
|||||
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaatcttccagcgactgatcgatatactgggccaactttgcccgcacaggggatccc 1560
Db 1660 AAAATCTTGGCCAGCGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagaccccaaggcccccaaatggcccccgctacacagcgggggtcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCCAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagctggacctgcgccgctggaggtgcgpggggggtgcgccccagggcctgc 1680
Db 1780 TACGTTAGCTGGACCTGCGGCCGCTGGAGGTGCGGGGGGCTGCGCGCCAGGCCTGC 1839
QY 1681 gccttctggaaccgtcttctcccccaaatgctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCCTCCCCAAATGCTCAGCGCCACC 1881

RESULT 10
US-08-370-156-3
; Sequence 3, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
US-08-370-156-3

Query Match 99.8%; Score 1722; DB 2; Length 3096;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagcccccgagtgctgctgcacacgccttccctggcttccccactccttctctc 60
Db 160 ATGAGCCCCCGCAGTGCTGCTGCACACGCGCTTCCTGGCTTCCCCCACTCCTTCTCTC 219
```

```
QY 61 ctctctggctcctgggtggaggagtgggggctgagggccgggagagatgcagagctgctg 120
Db 220 CTCCTCTGGCTCCTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279
QY 121 gtgacggtgcgtggggccggctcgggggcattcgccctgaagaccccccggggcccctgtc 180
Db 280 GTGACGGTGCCTGGGGCCGGCTCGGGGCAATTGCGCTGAAGACCCCCGGGGCCCTGTC 339
QY 181 tctgctttctgggcatcccccttttgcggagccccccatgggaccccgctcgtcttcttcca 240
Db 340 TCTGCTTTCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGTTCGCTTTCTGCCA 399
QY 241 ccggagcccaagcagccttggtcaggggtggtgtagacgctacaaaccttccagagtctgc 300
Db 400 CCGGAGCCCAAGCAGCCTTGGTCAGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacacccctatatccccagggttttgaggccagagatgtggaaccccaac 360
Db 460 TACCAATATGTGGACACCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcagagactgcctgtacctcaacgtgtggacacccataccccccggcctaca 420
Db 520 CGTGAGCTGACGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCCGGCCTACA 579
QY 421 tccccacccccctgctcctgctggtgatctatgtgggtgggtgttctacagtggggcctctcc 480
Db 580 TCCCCCACCCCTGTCTCCTGCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 639
QY 481 ttggacgtgtacgatggcgcgttcttggtagagccgagagactgtggtggtcctcatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 699
QY 541 aactaccgggtgggagccttggcttcttggcttccctgcccctgcccgggagccgagagccccgggc 600
Db 700 AACTACCGGTGGGAGCCTTTGGCTTCCTGGCCCTGCCGGGAGCCGAGAGGCCCGGGGC 759
QY 601 aatgtgggtcctcctggatcagaggctgcccctgagtggtgagagaaacgtggcagcc 660
Db 760 AATGTGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttccgggggtgacccccacatcagtgacgctgtttggggagagcgcgggagccgcctcgtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGAGCGCCCTCGGTG 879
QY 721 ggcattgcacctgctgtccccgcagcccgcccggttccacagggccgctgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCCGCCACGCGGGGCCCTGTTCCACAGGGCCCTGCTGCAGAGC 939
QY 781 ggtgcccccaatggaccctgggccagcgtgggcatgggagagcccgctcgagggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGCTGGGCGCATGGGAGAGGCCCGCTCGCAGGGCCACG 999
QY 841 cagctggccacacctgtgtggctgtcctccagcgcgccactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCACACCTTGTGGGCTGTCTCCAGGGCGGCACTGGTGGGAATGACACAGAGCTG 1059
QY 901 gtagcctgccttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
Db 1060 GTAGCCTGCCCTTCGGACACGACCACCGCAGGTCCTGGTGAACCCACGAATGGCACGTGCTG 1119
QY 961 cctcaagaaagcgtcttccgggttctccttccctgctgctggtgtagatggagacttccctcagt 1020
Db 1120 CCTCAAGAAAGCGCTTCCGGTTCTCCTTCGCTGCTGTGGTAGATGGAGACTTCTCTCAGT 1179
QY 1021 gacacccagagggccctcatcaacgctgggagacttccacggcctcaggtgctggtgggt 1080
Db 1180 GACACCCAGAGGCCCTCATCAACCGGGGAGACTTCCACGGCCTGCAGGTGTGTTGGTGGGT 1239
QY 1081 gtggtgaaggatgagggctcgtatttctgtgtttacggggccccagggcttcagcaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGTGTTTACGGGGCCCCCAGGGCTTCAGCAAAGAC 1299
```



```
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 27385..27387
; FEATURE:
; NAME/KEY: exon
; LOCATION: 28008..28129
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: terminator
; LOCATION: 28129..28131
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34528..34895)
; OTHER INFORMATION: /function= "arsenite resistance"
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (34092..34358)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33779..33963)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33493..33591)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (33297..33408)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32959..33094)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 6
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32569..32628)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (32386..32468)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31894..32080)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 9
; FEATURE:
```

```
; NAME/KEY: exon
; LOCATION: complement (31363..31534)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (31131..31284)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30816..31011)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30470..30626)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (30187..30274)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 14
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29945..30073)
; OTHER INFORMATION: /gene= "AR"
; OTHER INFORMATION: /number= 15
; FEATURE:
; NAME/KEY: exon
; LOCATION: complement (29664..29856)
; OTHER INFORMATION: /gene= "ARS"
; OTHER INFORMATION: /number= 16
; US-08-814-095-7

Query Match 62.0%; Score 1070; DB 3; Length 35060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1070; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgagccccgcagtgctgctgcacacgccttccctggcttccacactctcctc 60
   |||||||
Db 24110 ATGAGGCCCCGAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCTCTCCTC 24169

Qy 61 ctccctggtcctcctgggtggaggagtggtgggctgagggccggaggatgcagagctg 120
   |||||||
Db 24170 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGTGCTG 24229

Qy 121 gtgacggtcgctggggccgctgctgggggcattcgctgaagacccccggggccctgctc 180
   |||||||
Db 24230 GTGACGGTGGCTGGGGCCGCTGCGGGGCATTGCGCTGAAGACCCCGGGGCCCTGTC 24289

Qy 181 tctgcttctcctgggcatcccccttggcaggggtgtagacgctacaacctccagagtgctgc 240
   |||||||
Db 24290 TCTGCTTCTCTGGGCATCCCTTTGGGAGCCACCCCATGGGACCCCGCTGCTTCTGCCA 24349

Qy 241 ccggagcccaagcagccttggtcaggggtgtagacgctacaacctccagagtgctgc 300
   |||||||
Db 24350 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGCTGC 24409

Qy 301 taccaatatgtggacacccctataccagggttttggggcaccgagatgtggaaccccaac 360
   |||||||
Db 24410 TACCAATATGTGGACACCCCTATACCCAGGTTTGGGGCACCAGAGATGTGGAACCCCAAC 24469

Qy 361 cgtgagctgagcagggactgcctgtacctcaacgctgtggacacccataccccggcctaca 420
   |||||||
Db 24470 CGTGAGCTGAGCGGAGGACTGCCTGTACTCAACGTGTGGACACCATACCCCGGCCTACA 24529

Qy 421 tccccacacctgtcctcgtctggtatctatgggggtggtgcttctacagtggtggtcctcc 480
   |||||||
Db 24530 TCCCCACACCCCTGTCTCTCGTCTGGATCTATGGGGGTGGCTTCTTACAGTGGGGCCTCCTCC 24589
```

QY 481 ttggacgtgtacatggccgcttcttggtacagggccgagaggactgtgctggtgtccatg 540
|||||
Db 24590 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 24649
QY 541 aactacccgggtggagccctttggcttcttgccctgccggggagccgagagccccgggc 600
|||||
Db 24650 AACTACCGGGTGGAGCCTTTGGCTTCTTGCCCTGCCGGGAGCCGAGAGCCCCGGGC 24709
QY 601 aatgtgggtctctggatcagaggtggccctgcagtgggtgcaggagaacgtggcagcc 660
|||||
Db 24710 AATGTGGGTCTCTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCC 24769
QY 661 ttccgggggtgaccgacatcagtgacgtctgttggggagagcgcgggagccctcggtg 720
|||||
Db 24770 TTCGGGGGTACCGACATCAGTGACGCTGTGGGAGAGCGCGGGAGCGCCTCGGTG 24829
QY 721 ggcatacacctgtgtcccccgcagccggggcctgttccacagggccgtgctgcagagc 780
|||||
Db 24830 GGCATGCACCTGTCTCCCGCCAGCCGGGCGCTGTTCCACAGGGCGCTGCTGCAGAGC 24889
QY 781 ggtgcccccaatggacccttgggccacggtgggcatgggagagggccctgcagggccacg 840
|||||
Db 24890 GGTGCCCCCAATGCACCTTGGCCACCGTGGGCATGGGAGAGGCCCGTCCAGGGCCACG 24949
QY 841 cagctggccacattgtggctgtctccagggcgcactggtgggaatgacacagagctg 900
|||||
Db 24950 CAGCTGGCCACCTGTGGGCTGTCTCCAGCGGCACACTGGTGGGAATGACACAGAGCTG 25009
QY 901 gtacgtgcttgggacacgacgacgagcaggtctctggtgaaccacgacgtgctg 960
|||||
Db 25010 GTAGCCTGCCCTCGGACACGACGACGCGCAGGTCTGTTGAACCAACGATGGCAGCTG 25069
QY 961 cctcaagaaagcttccggttctctctgctgctgctgctgctggtgtagatggagacttcc 1020
|||||
Db 25070 CCTCAAGAAAGCTTCTCCGTTCTCCTTCGTCCTGTGGTAGATGGAGACTTCTCTCAGT 25129
QY 1021 gacacccagagccctcatcaacgcgggagacttccacggcctgcaggt 1070
|||||
Db 25130 GACACCCAGAGCCCTCATCAACGCGGGGAGACTTCCACGGCCTGCAGGT 25179

RESULT 13
US-08-370-156-13
; Sequence 13, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-370-156-23

Query Match 2.0%; Score 35; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1604 cggcggggggtcagcagtcagtagttagtctggacctg 1638
|||||
Db 35 CGCGGGGGGCTCAGCAGTACGTTAGTCTGGACCTG 1

RESULT 15

; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-370-156-13

Query Match 2.0%; Score 35; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1363 cgggtctacgctacgtcttgaacacccgtgcttc 1397
|||||
Db 1 CGGGTCTACGCCTACGTCTTTGAACACCCGTGCTTC 35

RESULT 14
US-08-370-156-23/c
; Sequence 23, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-370-156-23

Query Match 2.0%; Score 35; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1604 cggcggggggtcagcagtcagtagttagtctggacctg 1638
|||||
Db 35 CGCGGGGGGCTCAGCAGTACGTTAGTCTGGACCTG 1

RESULT 15

```
US-08-875-710-4
; Sequence 4, Application US/08875710
; Patent No. 6326139
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: zakut, Haim
; TITLE OF INVENTION: METHOD OF SCREENING FOR GENETIC PREDISPOSITION TO
; TITLE OF INVENTION: ANTICHOLINESTERASE THERAPY
; FILE REFERENCE: 2391.00076
; CURRENT APPLICATION NUMBER: US/08/875,710
; CURRENT FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: PCT/US96/00322
; EARLIER FILING DATE: 1996-01-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 35
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-875-710-4

Query Match      2.0%; Score 35; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1363 cgggtctacgcctacgtctttgaacaccgctgcttc 1397
      |||||||||||||||||||||||||||||||
Db 1 cgggtctacgcctacgtctttgaacaccgctgcttc 35

Search completed: August 31, 2002, 22:45:22
Job time: 19786 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:51:31 ; Search time 824.51 Seconds
(without alignments)
12008.885 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcattgcctgcaggtc.....ctatgaccatgattacgcc 5767

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5767	100.0	5767	24	Plasmid pTM034. S
2	2637	45.7	3858	22	Plasmid pRK50 used
3	2637	45.7	4960	22	PCMV-I-Cre-pa vect
4	2637	45.7	7332	22	PCMV-I-beta-pa vec
5	2637	45.7	8062	22	Plasmid pRK76 used
6	2637	45.7	8153	22	Plasmid pRK74 used
7	2637	45.7	8854	20	DNA encoding chime
8	2630	45.6	6295	19	Plasmid pWRG3196 e
9	2630	45.6	7164	22	Commercial plasmid

C 10	2630	45.6	7383	22	AAS00153	Matrix metalloprot
C 11	2628	45.6	4950	18	AAV03801	Retroviral vector
C 12	2628	45.6	4950	22	AAC82936	Transdominant effe
C 13	2628	45.6	5733	21	AA95416	Aspergillus nidula
C 14	2628	45.6	9164	20	AA82259	Beta-domain delete
C 15	2628	45.6	9632	20	AA04251	Baculovirus transf
C 16	2628	45.6	9780	20	AA82817	Plasmid pFLAG:BAC.
C 17	2628	45.6	11846	20	AA82261	Factor VIII protei
C 18	2628	45.6	12022	20	AA82260	Factor VIII protei
C 19	2586	44.8	5711	22	AAD04935	C31-Int expression
C 20	2577	44.7	4163	13	AAQ32349	Template clone Fab
C 21	2534	43.9	5897	20	AAV63741	Plasmid pTGP190-1
C 22	2534	43.9	6898	20	AAV63742	Plasmid UGP232-4 c
C 23	2534	43.9	9335	20	AAV63734	Plasmid PIGP/367 c
C 24	2533	43.9	3753	24	ABA04129	Plasmid pHCE19T(II
C 25	2533	43.9	3755	24	ABA04130	Plasmid pHCE19T(II
C 26	2533	43.9	4283	18	AAT86449	DNA encoding hGH.
C 27	2533	43.9	4283	19	AAT86610	Epidermal growth f
C 28	2533	43.9	4283	21	AAA63237	Plasmid pWRG1630 n
C 29	2533	43.9	6565	22	AAF55126	Nucleotide sequenc
C 30	2533	43.9	6971	22	AAF55124	Nucleotide sequenc
C 31	2533	43.9	7558	22	AAF55125	Nucleotide sequenc
C 32	2533	43.9	7969	22	AAF55123	Nucleotide sequenc
C 33	2528	43.8	4776	20	AA77617	Expression constru
C 34	2528	43.8	4776	20	AA77614	Expression constru
C 35	2527	43.8	7566	14	AAQ42160	Plasmid pPS0212 co
C 36	2527	43.8	7639	14	AAQ42159	Plasmid pJD884 con
C 37	2526	43.8	4713	19	AAV12067	Murine IAD alpha c
C 38	2526	43.8	4724	19	AAV12068	Murine IAD beta ch
C 39	2526	43.8	5421	21	AAC68299	SV40/APPA plasmid
C 40	2526	43.8	5465	20	AA220088	Plasmid PUBIAC. C
C 41	2526	43.8	6116	21	AAC68297	R15/APPA plasmid c
C 42	2526	43.8	6708	21	AAC68295	R15/APPA plasmid c
C 43	2526	43.8	9534	18	AAT62072	Vector RP3224E2 en
C 44	2509	43.5	7380	20	AA84028	MMP9 promoter beta
C 45	2480	43.0	4045	15	AAQ70942	Plasmid pND211 con

ALIGNMENTS

RESULT 1	
AAS17547	
ID	AAS17547 standard; DNA; 5767 BP.
XX	
AC	AAS17547;
XX	
DT	25-FEB-2002 (first entry)
XX	
DE	Plasmid pTM034.
XX	
KW	Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW	transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW	muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW	organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW	circular; ds.
OS	Synthetic.
XX	
PN	WO200171014-A2.
PD	27-SEP-2001.
XX	
PF	16-MAR-2001; 2001WO-US08468.
XX	
PR	17-MAR-2000; 2000US-190440P.
XX	
PA	(MORT/) MOR T.
PA	(SORE/) SOREQ H.
PA	(ARNT/) ARNTZEN C.
PA	(MASO/) MASON H.
PA	(BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX	

PI Mor T, Soreq H, Arntzen C, Mason H;

XX WPI; 2002-055120/07.

XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -

XX Claim 11; Page 28-31; 42pp; English.

XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM034, the pGPRVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.

XX
SQ Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;

Query Match 100.0%; Score 5767; DB 24; Length 5767;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	agcttgcatgcctgcaggtcaacatggtggagcagcacactctcgtctactccaagaata	60
Db	1	agcttgcatgcctgcaggtcaacatggtggagcagcacactctcgtctactccaagaata	60
QY	61	tcaagatacagtcctcagaagaccagagggctattgagacttttcaacaaagggtaatat	120
Db	61	tcaagatacagtcctcagaagaccagagggctattgagacttttcaacaaagggtaatat	120
QY	121	cgggaaacctctcggattccattgccagctatctgtcacttcacgaagggacagtag	180
Db	121	cgggaaacctctcggattccattgccagctatctgtcacttcacgaagggacagtag	180
QY	181	aaaaggaagatggcttctacaaatgccatcattgcgataaaaggaaaggctatcgttcaag	240
Db	181	aaaaggaagatggcttctacaaatgccatcattgcgataaaaggaaaggctatcgttcaag	240
QY	241	aatgcctctaccgacagtggtcccaagatggacccccaccacaggaacatcgtggaa	300
Db	241	aatgcctctaccgacagtggtcccaagatggacccccaccacaggaacatcgtggaa	300
QY	301	aaagaagagcttccaaccacgctcttcaagcaagtggattgatgtgataacttttcaaca	360
Db	301	aaagaagagcttccaaccacgctcttcaagcaagtggattgatgtgataacttttcaaca	360
QY	361	aagggtaatatcgggaaacctcctcgattccattgccagctatctgtcacttcacga	420
Db	361	aagggtaatatcgggaaacctcctcgattccattgccagctatctgtcacttcacga	420
QY	421	aaggacagtagaaaaaggaagatggcttctacaaatgccatcattgcgataaaggaaaggc	480
Db	421	aaggacagtagaaaaaggaagatggcttctacaaatgccatcattgcgataaaggaaaggc	480

QY	481	tatcgttcaagaatgcctctaccgacagtggtcccaagatggacccccaccacagagga	540
Db	481	tatcgttcaagaatgcctctaccgacagtggtcccaagatggacccccaccacagagga	540
QY	541	acatcgtgaaaaaagaagacggttccaaccacgctcttcaaagcaagtgtgattgtgata	600
Db	541	acatcgtgaaaaaagaagacggttccaaccacgctcttcaaagcaagtgtgattgtgata	600
QY	601	tctccactgacgtaaggatgacgcacaatccccactatccttcgcgaagacccttccctta	660
Db	601	tctccactgacgtaaggatgacgcacaatccccactatccttcgcgaagacccttccctta	660
QY	661	tataaggaagtcatctcatttgcatttgagagagacctcgagaatttcaacacaacatat	720
Db	661	tataaggaagtcatctcatttgcatttgagagagacctcgagaatttcaacacaacatat	720
QY	721	acaaaacaaacgaatctcaagcaatcttctgaaattttcaccatttacgaatttaaatca	780
Db	721	acaaaacaaacgaatctcaagcaatcttctgaaattttcaccatttacgaatttaaatca	780
QY	781	ttcttttaagcaaaagcaattttctgaaattttcaccatttacgaatttaaatca	840
Db	781	ttcttttaagcaaaagcaattttctgaaattttcaccatttacgaatttaaatca	840
QY	841	gctccccgcagtgctgctgcacacgccttccctggcttccccactccttctcctctc	900
Db	841	gctccccgcagtgctgctgcacacgccttccctggcttccccactccttctcctctc	900
QY	901	ctctggctcctgggtggagagtggtgggctgagggccggaggatgcagagctgctggtg	960
Db	901	ctctggctcctgggtggagagtggtgggctgagggccggaggatgcagagctgctggtg	960
QY	961	acggtgctggggggccggtgcggtgggcatctgcctgaagacccccggggccctgtctct	1020
Db	961	acggtgctggggggccggtgcggtgggcatctgcctgaagacccccggggccctgtctct	1020
QY	1021	gctttcctgggcatcccccttgcggagccacccatgggacccccgtgctttctgccaccg	1080
Db	1021	gctttcctgggcatcccccttgcggagccacccatgggacccccgtgctttctgccaccg	1080
QY	1081	gagcccaagcagccttggtcaggggtggtagacgctacaacaccttccagagtgtctgtac	1140
Db	1081	gagcccaagcagccttggtcaggggtggtagacgctacaacaccttccagagtgtctgtac	1140
QY	1141	caatatgtgacacccctatataccagggttttgagggccaccgagatgtggaacccccaccgt	1200
Db	1141	caatatgtgacacccctatataccagggttttgagggccaccgagatgtggaacccccaccgt	1200
QY	1201	gagctgagcgaggactgcctgtacctcaacgctgtggacaccataccccggcctacatcc	1260
Db	1201	gagctgagcgaggactgcctgtacctcaacgctgtggacaccataccccggcctacatcc	1260
QY	1261	cccacccctgctcctcgtcgtggtatctatgggggtggtcttctacagtgggccctcctctg	1320
Db	1261	cccacccctgctcctcgtcgtggtatctatgggggtggtcttctacagtgggccctcctctg	1320
QY	1321	gacgtgtacgatggcgccttcttgcgtacagggccgagaggactgtgctggtgccatgaac	1380
Db	1321	gacgtgtacgatggcgccttcttgcgtacagggccgagaggactgtgctggtgccatgaac	1380
QY	1381	taccgggtggagaccccttggcttccctggcctgcggggagccgagagggcccggaat	1440
Db	1381	taccgggtggagaccccttggcttccctggcctgcggggagccgagagggcccggaat	1440
QY	1441	gtgggtctcctggatcagaggctggccctgcagtggtgagcaggagaacgtggcagccttc	1500
Db	1441	gtgggtctcctggatcagaggctggccctgcagtggtgagcaggagaacgtggcagccttc	1500
QY	1501	gggggtgacccgacatcagtgacgctgtttggggagagcgcgggagccctcggtgggc	1560
Db	1501	gggggtgacccgacatcagtgacgctgtttggggagagcgcgggagccctcggtgggc	1560

QY 1561 atgcacctgctgtccccgccagcccggggcctgttccacagggccggtgctgcagagcggt 1620
Db 1561 atgcacctgctgtccccgccagcccggggcctgttccacagggccggtgctgcagagcggt 1620
QY 1621 gcccccaatggaccctgggccacggtgggcatgggagagggcccgctgcagggccacgcag 1680
Db 1621 gcccccaatggaccctgggccacggtgggcatgggagagggcccgctgcagggccacgcag 1680
QY 1681 ctggcccaacctgtgggctgtccctccaggcggcactggtggaatgacacagagctggta 1740
Db 1681 ctggcccaacctgtgggctgtccctccaggcggcactggtggaatgacacagagctggta 1740
QY 1741 gccctgccttcggacacgacccagcgaggtccctggtgaacacagaaatggcacgtgctgcct 1800
Db 1741 gccctgccttcggacacgacccagcgaggtccctggtgaacacagaaatggcacgtgctgcct 1800
QY 1801 caagaaagcgtcttcccggttctccctcgtgcctgctggtgtagatggagacttcctcagtgac 1860
Db 1801 caagaaagcgtcttcccggttctccctcgtgcctgctggtgtagatggagacttcctcagtgac 1860
QY 1861 accccagagggcctcatcaacggggagagacttccacggcctgcaggtgctggtggtg 1920
Db 1861 accccagagggcctcatcaacggggagagacttccacggcctgcaggtgctggtggtg 1920
QY 1921 gtgaaggatgagggctcgtatttctggttttacggggcccccaggcttcagcaaaagacaac 1980
Db 1921 gtgaaggatgagggctcgtatttctggttttacggggcccccaggcttcagcaaaagacaac 1980
QY 1981 gagtctctcatcagccgggcgaggtccctggccggggtcggggtcggggttccccaggta 2040
Db 1981 gagtctctcatcagccgggcgaggttccctggccggggtcggggtcggggttccccaggta 2040
QY 2041 agtgacctggcagccgaggtggtcctgcattacacagactggctgcacccccaggac 2100
Db 2041 agtgacctggcagccgaggtggtcctgcattacacagactggctgcacccccaggac 2100
QY 2101 ccggcacgcctgagggagggccctgagcgtggtggtggggccaccacaatgtcgtgtgcccc 2160
Db 2101 ccggcacgcctgagggagggccctgagcgtggtggtggggccaccacaatgtcgtgtgcccc 2160
QY 2161 gtggccagctggctggggcgactggctgcccagggtggcccggtctacgcctactcttt 2220
Db 2161 gtggccagctggctggggcgactggctgcccagggtggcccggtctacgcctactcttt 2220
QY 2221 gaacacgctgcttccacgctctcctgcccctgtggatgggggtgccccagggctacgag 2280
Db 2221 gaacacgctgcttccacgctctcctgcccctgtggatgggggtgccccagggctacgag 2280
QY 2281 atcgagttcatcttgggataccccctggaccctctcgaaactacacggcagaggagaaa 2340
Db 2281 atcgagttcatcttgggataccccctggaccctctcgaaactacacggcagaggagaaa 2340
QY 2341 atcttgcgccagcgactgatgcgatactgggccaaacttggccgaacttggccgcacagggatcccaat 2400
Db 2341 atcttgcgccagcgactgatgcgatactgggccaaacttggccgaacttggccgcacagggatcccaat 2400
QY 2401 gagccccgagaccccaaggccccacaatggcccccgtaacggcgggggctcagcagttac 2460
Db 2401 gagccccgagaccccaaggccccacaatggcccccgtaacggcgggggctcagcagttac 2460
QY 2461 gttagttctggacctgcgccgctggagggtgcggcgggggtgcgcgccccagcctgcgcc 2520
Db 2461 gttagttctggacctgcgccgctggagggtgcggcgggggtgcggcgccccagcctgcgcc 2520
QY 2521 ttctggaaaccgcttccctccccaaaattgctcagcgctacctcgtataggtaccgagctctctc 2580
Db 2521 ttctggaaaccgcttccctccccaaaattgctcagcgctacctcgtataggtaccgagctctctc 2580
QY 2581 aacaatctagctagagtttgcctctatctatatgtataaagggtatgctgatatgcactat 2640
Db 2581 aacaatctagctagagtttgcctctatctatatgtataaagggtatgctgatatgcactat 2640
QY 2641 tcaaataggagcattagctatgttgttaattgtcacttttatgttatgtgggtaagtcacc 2700

Db 2641 tcaaataggagcattagctatgttgttaattgtcactttatgttatgtgggtaagtcacc 2700
QY 2701 taagacactccacgtacacctacgtttgtgtctcttaccggctttaataaatcttctgcct 2760
Db 2701 taagacactccacgtacacctacgtttgtgtctcttaccggctttaataaatcttctgcct 2760
QY 2761 tgttccataatttactaattatcccttcttcaactaaaagaaaaatgttatcattaaagtat 2820
Db 2761 tgttccataatttactaattatcccttcttcaactaaaagaaaaatgttatcattaaagtat 2820
QY 2821 tagtctttagaacatatagaggctctttaattgggtagggttttacaanaataactataaaa 2880
Db 2821 tagtctttagaacatatagaggctctttaattgggtagggttttacaanaataactataaaa 2880
QY 2881 atgtcataaaaaatccacgtgggttaaaaaaaatgcagaaaaatcgacgtcgtctattggaccga 2940
Db 2881 atgtcataaaaaatccacgtgggttaaaaaaaatgcagaaaaatcgacgtcgtctattggaccga 2940
QY 2941 cagttgctattataataataatggggccacatatagtagctgacaaaataaaattacctgacaaca 3000
Db 2941 cagttgctattataataatggggccacatatagtagctgacaaaataaaattacctgacaaca 3000
QY 3001 tcgtttcactaaaataacaaacacaaaaaggagtgcattttccagggcatttttgttaata 3060
Db 3001 tcgtttcactaaaataacaaacacaaaaaggagtgcattttccagggcatttttgttaata 3060
QY 3061 aaaaacagttaaaaggagtgcaatagaaaaatatagggtgtggaatatgtatttgagca 3120
Db 3061 aaaaacagttaaaaggagtgcaatagaaaaatatagggtgtggaatatgtatttgagca 3120
QY 3121 cgtcttgaagcgaaatttcactggccgtcgttttacaacgctcgtgactgggaaaaaccctggc 3180
Db 3121 cgtcttgaagcgaaatttcactggccgtcgttttacaacgctcgtgactgggaaaaaccctggc 3180
QY 3181 gttacccaaacttaatcgcccttcccaacagatcccccttctgccagctggcgtataagcgaa 3240
Db 3181 gttacccaaacttaatcgcccttcccaacagatcccccttctgccagctggcgtataagcgaa 3240
QY 3241 gagggccgcacccgatcgcccttcccaacagttgcgcagcctgaatggcgaaatggcgctg 3300
Db 3241 gagggccgcacccgatcgcccttcccaacagttgcgcagcctgaatggcgaaatggcgctg 3300
QY 3301 atgcggtattttctccttacgcatactgtgcgggtatttcacacccgcataatggtgcactctc 3360
Db 3301 atgcggtattttctccttacgcatactgtgcgggtatttcacacccgcataatggtgcactctc 3360
QY 3361 agtacaaatctgctctgatgcgcgcataagtaagccagccccgcagacacccgcacacccgct 3420
Db 3361 agtacaaatctgctctgatgcgcgcataagtaagccagccccgcagacacccgcacacccgct 3420
QY 3421 gacgcgccctgacgggcttgcgtcgtcccgccgcataccgcttacagacaagctgtgaccgtc 3480
Db 3421 gacgcgccctgacgggcttgcgtcgtcccgccgcataccgcttacagacaagctgtgaccgtc 3480
QY 3481 tccggggagctgcatgtgtcagaggttttcacccgtcatcacccgaaacgcgcgagacgaaaag 3540
Db 3481 tccggggagctgcatgtgtcagaggttttcacccgtcatcacccgaaacgcgcgagacgaaaag 3540
QY 3541 ggcctcgtgatacgccctatttttatagggttaattgcatgataaataatggtttcttagacg 3600
Db 3541 ggcctcgtgatacgccctatttttatagggttaattgcatgataaataatggtttcttagacg 3600
QY 3601 tcagggtggcaccttttcgggggaaaatgtgcgcggaaacccctatttgtttatttttctaaata 3660
Db 3601 tcagggtggcaccttttcgggggaaaatgtgcgcggaaacccctatttgtttatttttctaaata 3660
QY 3661 cattcaaatatgtatccgcctcatggagacaataaacctgatataatcttcaataatatga 3720
Db 3661 cattcaaatatgtatccgcctcatggagacaataaacctgatataatcttcaataatatga 3720
QY 3721 aaaagggaagagtagtagtattcaacatttcogtgcgcccttatcccttttttgcgggca 3780
Db 3721 aaaagggaagagtagtagtattcaacatttcogtgcgcccttatcccttttttgcgggca 3780

Db 3721 aaaaggaagagtatgagtattcaacatttcctggtgcgccctattcccttttttgcggca 3780

Qy 3781 ttttgccttcctgtttttgctcacccagaaacgctggtgaaaagtaaaagatgctgaagat 3840

Db 3781 ttttgccttcctgtttttgctcacccagaaacgctggtgaaaagtaaaagatgctgaagat 3840

Qy 3841 cagttgggtgcacgagtggtgttacatgcgaactggatctcaacagcggtaagatccttgag 3900

Db 3841 cagttgggtgcacgagtggtgttacatgcgaactggatctcaacagcggtaagatccttgag 3900

Qy 3901 agttttcccccgaagaacggttttccaatgatgagcacatttttaaagttctgctatggtc 3960

Db 3901 agttttcccccgaagaacggttttccaatgatgagcacatttttaaagttctgctatggtc 3960

Qy 3961 gcggtattatcccgatttgacgccgggcaagagcaactcggtcgcgcgcatacactattct 4020

Db 3961 gcggtattatcccgatttgacgccgggcaagagcaactcggtcgcgcgcatacactattct 4020

Qy 4021 cagaatgacttggttgagtactcacagaaaaagcatcttacggatggcatgaca 4080

Db 4021 cagaatgacttggttgagtactcacagaaaaagcatcttacggatggcatgaca 4080

Qy 4081 gtaagagaattatgcagtgctgccataaccatgagtgataaacactgcggccaaacttactt 4140

Db 4081 gtaagagaattatgcagtgctgccataaccatgagtgataaacactgcggccaaacttactt 4140

Qy 4141 ctgacaaacgatcggaggaccgaaggaagctaacccgttttttgcacaaacatggggatcat 4200

Db 4141 ctgacaaacgatcggaggaccgaaggaagctaacccgttttttgcacaaacatggggatcat 4200

Qy 4201 gtaactgccttgatcgttgggaacgggagctgaatgaagccataccaaacgacgagcgt 4260

Db 4201 gtaactgccttgatcgttgggaacgggagctgaatgaagccataccaaacgacgagcgt 4260

Qy 4261 gacaccagatgcctgtagcaattggcaacaacgttgcgcaaaactattaaactggcgaacta 4320

Db 4261 gacaccagatgcctgtagcaattggcaacaacgttgcgcaaaactattaaactggcgaacta 4320

Qy 4321 cttactctagcttcccggcaacaataatagactggatggaggcggataaaagtgcagga 4380

Db 4321 cttactctagcttcccggcaacaataatagactggatggaggcggataaaagtgcagga 4380

Qy 4381 ccacttctgcgtcgcgccttccggctggctggtttattgctgataaaatctggagccggt 4440

Db 4381 ccacttctgcgtcgcgccttccggctggctggtttattgctgataaaatctggagccggt 4440

Qy 4441 gagcgtgggtctcgcggtatcatctgagcactggggccagatggtaaagccctcccgtatc 4500

Db 4441 gagcgtgggtctcgcggtatcatctgagcactggggccagatggtaaagccctcccgtatc 4500

Qy 4501 gtagttatctacacgacggtggatgaaggcaactatggatgaacgaaatagacagatcgct 4560

Db 4501 gtagttatctacacgacggtggatgaaggcaactatggatgaacgaaatagacagatcgct 4560

Qy 4561 gagataggtgcctcactgattaaagcattggtaactgtcagaccacgaagtttactcatata 4620

Db 4561 gagataggtgcctcactgattaaagcattggtaactgtcagaccacgaagtttactcatata 4620

Qy 4621 ctttagattgatttaaaaacttcaatttttaatttaaaggatctaggtgaagatcccttttt 4680

Db 4621 ctttagattgatttaaaaacttcaatttttaatttaaaggatctaggtgaagatcccttttt 4680

Qy 4681 gataatctcatgaccacaaatcccttaacgtgagtttcttcgttccactgagcgtcagacccc 4740

Db 4681 gataatctcatgaccacaaatcccttaacgtgagtttcttcgttccactgagcgtcagacccc 4740

Qy 4741 gtagaaaagatcaaaaggatcttcttgagatcccttttttctgcggtaatctgctgcttg 4800

Db 4741 gtagaaaagatcaaaaggatcttcttgagatcccttttttctgcggtaatctgctgcttg 4800

Qy 4801 caaacacacacaccccgctaccagcgggtggtttgtttgcccggatcaagagctaccaact 4860

Db 4801 caaacacacacacaccccgctaccagcgggtggtttgtttgcccggatcaagagctaccaact 4860

Qy 4861 cttttccgaaggtaaactggcttcagcagcgcagataccaaaatactgtccttctagt 4920

Db 4861 cttttccgaaggtaaactggcttcagcagcgcagataccaaaatactgtccttctagt 4920

Qy 4921 tagccgtagttaggccaccacttcaagaactctgttagcaccgcctacatacctcgtctg 4980

Db 4921 tagccgtagttaggccaccacttcaagaactctgttagcaccgcctacatacctcgtctg 4980

Qy 4981 ctaatcctgttacccagtggtgctgccagtgggcgataaagtcgtctctaccgggttgac 5040

Db 4981 ctaatcctgttacccagtggtgctgccagtgggcgataaagtcgtctctaccgggttgac 5040

Qy 5041 tcaagacgataagttaccgggataaaggcgcagcgggtcgggctgaacgggggttcgtgcaca 5100

Db 5041 tcaagacgataagttaccgggataaaggcgcagcgggtcgggctgaacgggggttcgtgcaca 5100

Qy 5101 cagcccagcttgagcgaaacgacctacacgaaactgagatacctacagcgtgagctatga 5160

Db 5101 cagcccagcttgagcgaaacgacctacacgaaactgagatacctacagcgtgagctatga 5160

Qy 5161 gaaagcgccacgcttcccgaaggagaaaggcggaaggatatccggtaaagcggcagggctc 5220

Db 5161 gaaagcgccacgcttcccgaaggagaaaggcggaaggatatccggtaaagcggcagggctc 5220

Qy 5221 ggaacagaggagcgcacgagggagccttccaggggggaaacgcctggtatctttatagtcct 5280

Db 5221 ggaacagaggagcgcacgagggagccttccaggggggaaacgcctggtatctttatagtcct 5280

Qy 5281 gtcggggttccgccacctctgacttgagcgtcgatttttgtgatgctgcaggggggcgg 5340

Db 5281 gtcggggttccgccacctctgacttgagcgtcgatttttgtgatgctgcaggggggcgg 5340

Qy 5341 agcctatggaaaaaacgccagcaacgggccttttttacggttccctggccttttgctggcct 5400

Db 5341 agcctatggaaaaaacgccagcaacgggccttttttacggttccctggccttttgctggcct 5400

Qy 5401 tttgctcacatgttcttctcgtggttatccctgattctgtggataaacgctattaccgcc 5460

Db 5401 tttgctcacatgttcttctcgtggttatccctgattctgtggataaacgctattaccgcc 5460

Qy 5461 tttgagtgagctgataccgctcgcgcgcagcgaacgacgagcgcgagcgtcagtgagc 5520

Db 5461 tttgagtgagctgataccgctcgcgcgcagcgaacgacgagcgcgagcgtcagtgagc 5520

Qy 5521 gaggaagcggaaagagcgcccaatacgcgcaaacgcctctcccgcgcttgccgattcat 5580

Db 5521 gaggaagcggaaagagcgcccaatacgcgcaaacgcctctcccgcgcttgccgattcat 5580

Qy 5581 taatgcagctggcagcagaggtttcccgactgaaagcgggagtgagcgaacgcaatt 5640

Db 5581 taatgcagctggcagcagaggtttcccgactgaaagcgggagtgagcgaacgcaatt 5640

Qy 5641 aatgtgagttagctcaactcattagggaccaccaggctttacactttatgcttccggctcgt 5700

Db 5641 aatgtgagttagctcaactcattagggaccaccaggctttacactttatgcttccggctcgt 5700

Qy 5701 atgttgtgtggaattgtgagcgggataaacaatttcacacaggaaacagctatgacctgat 5760

Db 5701 atgttgtgtggaattgtgagcgggataaacaatttcacacaggaaacagctatgacctgat 5760

Qy 5761 tacgcca 5767

Db 5761 tacgcca 5767

RESULT 2
AAD04947
ID AAD04947 standard; DNA; 3858 BP.
XX
AC AAD04947;
XX
DT 17-JUL-2001 (first entry)

Db 2349 ttaatcgcttgagcacatcccccttctcgccagctggcgtaataagcgaaggcccgca 2408
QY 3251 ccgatacgcccttcccaacagttgcgagccctgaatggcgcaatggcgccctgatgcggtatt 3310
Db 2409 ccgatacgcccttcccaacagttgcgagccctgaatggcgcaatggcgccctgatgcggtatt 2468
QY 3311 ttctccttacgcatactgtgcggtatttcacaccgcataatggtgcactctcagtaacaatc 3370
Db 2469 ttctccttacgcatactgtgcggtatttcacaccgcataatggtgcactctcagtaacaatc 2528
QY 3371 gctctgatgccgcatagtttaagccagccccgacacccccgcaacccccgctgacgcgccct 3430
Db 2529 gctctgatgccgcatagtttaagccagccccgacacccccgcaacccccgctgacgcgccct 2588
QY 3431 gacgggcttgctgctgctcccgcatccgcttacagacaaagctgtgaccgctctccgggagct 3490
Db 2589 gacgggcttgctgctgctcccgcatccgcttacagacaaagctgtgaccgctctccgggagct 2648
QY 3491 gcatgtgtcagagggttttcaccgctcatcaccgaaacgcgcgagacgaaaggccctcgtga 3550
Db 2649 gcatgtgtcagagggttttcaccgctcatcaccgaaacgcgcgagacgaaaggccctcgtga 2708
QY 3551 taqgcctatttttatagggttaatgtcatgataataatggtttcttagacgtcaggtggca 3610
Db 2709 taqgcctatttttatagggttaatgtcatgataataatggtttcttagacgtcaggtggca 2768
QY 3611 cttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacatattcaata 3670
Db 2769 cttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacatattcaata 2828
QY 3671 tgtatccgctcatgagacaaataaccctgataaaatgcttcaataatatgaaaaaggaaaga 3730
Db 2829 tgtatccgctcatgagacaaataaccctgataaaatgcttcaataatatgaaaaaggaaaga 2888
QY 3731 gtagagatttcaacatttccggtgcgcccttattcccttttttgcggcattttgccttc 3790
Db 2889 gtagagatttcaacatttccggtgcgcccttattcccttttttgcggcattttgccttc 2948
QY 3791 ctgtttttgctcaccocagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 3850
Db 2949 ctgtttttgctcaccocagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 3008
QY 3851 cacagtggttacatcgaaactggatctcaacagcggtaagatccttgagagtttccgcc 3910
Db 3009 cacagtggttacatcgaaactggatctcaacagcggtaagatccttgagagtttccgcc 3068
QY 3911 ccgaagaacgttttccaaatgatgagcacattttaaagtcttgctatgtggcggtattat 3970
Db 3069 ccgaagaacgttttccaaatgatgagcacattttaaagtcttgctatgtggcggtattat 3128
QY 3971 cccgtattgacgccgggcaagcaactcggctgcgcgcatacacatttctcagaatgact 4030
Db 3129 cccgtattgacgccgggcaagagcaactcggctgcgcgcatacacatttctcagaatgact 3188
QY 4031 tgggtgagtactcaccagtcacagaaaaagcatctttacggatggcatgacagtaagagaat 4090
Db 3189 tgggtgagtactcaccagtcacagaaaaagcatctttacggatggcatgacagtaagagaat 3248
QY 4091 tatgcagtgctgccataaccaatgagtgataaacactgcggccaacttacttctgacaaacga 4150
Db 3249 tatgcagtgctgccataaccaatgagtgataaacactgcggccaacttacttctgacaaacga 3308
QY 4151 tcggaggaccgaaggagtaaccgcttttttgacaaacatgggggatcatgtaaactcgcc 4210
Db 3309 tcggaggaccgaaggagtaaccgcttttttgacaaacatgggggatcatgtaaactcgcc 3368
QY 4211 ttgatcgttgggaaccggagctgaatgaagccataccaaacagcagcgtgacaccacga 4270
Db 3369 ttgatcgttgggaaccggagctgaatgaagccataccaaacagcagcgtgacaccacga 3428
QY 4271 tgcctgtagcaatggcaacaacgcttgcgcaaacatttaactggcggaactacttactctag 4330
|||||

Db 3429 tgccctgtagcaatggcaacaacgttgcgcaaacattattaactggcgaaactacttactctag 3488
QY 4331 ctcccgggcaacaattaatagactggatggagcggataaaagttgcaggaccacttctgc 4390
Db 3489 ctcccgggcaacaattaatagactggatggagcggataaaagttgcaggaccacttctgc 3548
QY 4391 gctcgcccttccggctggtgtttattgctgataaaatctgagccggtgagcgtgggt 4450
Db 3549 gctcgcccttccggctggtgtttattgctgataaaatctgagccggtgagcgtgggt 3608
QY 4451 ctcgcggtatcatctgcagcaactggggccagatggtaaagccctcccgatatcgtagtattc 4510
Db 3609 ctcgcggtatcatctgcagcaactggggccagatggtaaagccctcccgatatcgtagtattc 3668
QY 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg 4570
Db 3669 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg 3728
QY 4571 cctcactgattaaagcattggtaactgtcagaccaccaagtttactcatatatactttagattg 4630
Db 3729 cctcactgattaaagcattggtaactgtcagaccaccaagtttactcatatatactttagattg 3788
QY 4631 atttaaaacttcattttttaatttaaaaggatctagggtgaagatccctttttgataaattca 4690
Db 3789 atttaaaacttcattttttaatttaaaaggatctagggtgaagatccctttttgataaattca 3848
QY 4691 tgaccaaaaatcccccttaacgtgagtttctgcgtccactgagcgtcagacccccgtagaaaaa 4750
Db 3849 tgaccaaaaatcccccttaacgtgagtttctgcgtccactgagcgtcagacccccgtagaaaaa 3908
QY 4751 tcaaggatcttcttgagatcccttttttctgcgcgtaactctgctgcttgcacacaaaaa 4810
Db 3909 tcaaggatcttcttgagatcccttttttctgcgcgtaactctgctgcttgcacacaaaaa 3968
QY 4811 aaccacgctaccagcggtgtgttttgcggatcaagagctaccacactcttttccga 4870
Db 3969 aaccacgctaccagcggtgtgttttgcggatcaagagctaccacactcttttccga 4028
QY 4871 aggtaaactggcttcagcagagcgcagataccaaaatactgtccttctagtgtagccgtagt 4930
Db 4029 aggtaaactggcttcagcagagcgcagataccaaaatactgtccttctagtgtagccgtagt 4088
QY 4931 tagggcaccacttcaagaactctgtaggacccgctacataacctcgctctgtaaatccctgt 4990
Db 4089 tagggcaccacttcaagaactctgtaggacccgctacataacctcgctctgtaaatccctgt 4148
QY 4991 taccagtggctgtgccagtgcgataaagtcgtgtcttaccgggttggaactcaagacgat 5050
Db 4149 taccagtggctgtgccagtgcgataaagtcgtgtcttaccgggttggaactcaagacgat 4208
QY 5051 agttaccgggataaggcgagcgggtcggtgaaacgggggttcgtgacacagcccagct 5110
Db 4209 agttaccgggataaggcgagcgggtcggtgaaacgggggttcgtgacacagcccagct 4268
QY 5111 tggagcgaacgacctacacccgaactgagataacctacagcgtgagctatgagaaagcgcca 5170
Db 4269 tggagcgaacgacctacacccgaactgagataacctacagcgtgagctatgagaaagcgcca 4328
QY 5171 cgcttcccgaaaggagaaaaaggcgacaggtatccggtaagcggcaggggtcggaacagggag 5230
Db 4329 cgcttcccgaaaggagaaaaaggcgacaggtatccggtaagcggcaggggtcggaacagggag 4388
QY 5231 agcgcaagaggagcttccaggggggaaacgcctggtatctttatagtcctgtcggttttc 5290
Db 4389 agcgcaagaggagcttccaggggggaaacgcctggtatctttatagtcctgtcggttttc 4448
QY 5291 gccacctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatgga 5350
Db 4449 gccacctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatgga 4508
QY 5351 aaaacgcagcaacgcggcccttttttacgggttcctggcccttttgccttttgcctcaca 5410
Db 4509 aaaacgcagcaacgcggcccttttttacgggttcctggcccttttgccttttgcctcaca 4568

Qy 5411 tgttcttctcgttataccccctgattctgttgataaaccgtattaccgcctttgagtga 5470
|||||
Db 4569 tgttcttctcgttataccccctgattctgttgataaaccgtattaccgcctttgagtga 4628
|||||
Qy 5471 ctgataccgctcgcgcagcagccgaacgagcgcagcgcagcgcagcgcagcgcagcgcg 5530
|||||
Db 4629 ctgataccgctcgcgcagcagccgaacgagcgcagcgcagcgcagcgcagcgcagcgcg 4688
|||||
Qy 5531 aagagcgcaccaatacgcgaacccctctctcccgcgcggttgccgattcattaatgcagct 5590
|||||
Db 4689 aagagcgcaccaatacgcgaacccctctctcccgcgcggttgccgattcattaatgcagct 4748
|||||
Qy 5591 ggcacgacaggtttcccgactggaagcgggcagtgagcgcgaacgcaaatattatgtgagtt 5650
|||||
Db 4749 ggcacgacaggtttcccgactggaagcgggcagtgagcgcgaacgcaaatattatgtgagtt 4808
|||||
Qy 5651 agctcactcaattagggcaccgccaggtttacactttatgcttccggctcgtatgtgtgtg 5710
|||||
Db 4809 agctcactcaattagggcaccgccaggtttacactttatgcttccggctcgtatgtgtgtg 4868
|||||
Qy 5711 gaattgtgagcgggataacaatttcacacaggaacagctatgaccatgattacgccca 5767
|||||
Db 4869 gaattgtgagcgggataacaatttcacacaggaacagctatgaccatgattacgccca 4925
|||||

RESULT 4

AAD09270
ID AAD09270 standard; DNA; 7332 BP.
XX
AC AAD09270;
XX
DT 12-SEP-2001 (first entry)
XX
DE pCMV-I-beta-pA vector DNA related to the invention.
XX
KW DNA recombinase domain; protein transduction domain; PTD;
KW gene alteration; fusion protein; Human immunodeficiency virus;
KW HIV; pCMV-I-beta-pA vector; ds.
XX
OS Unidentified.
XX
PN WO200149832-A2.
XX
PD 12-JUL-2001.
XX
PF 05-JAN-2001; 2001WO-EP000060.
XX
PR 07-JAN-2000; 2000EP-0100351.
PR 10-NOV-2000; 2000EP-0124595.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
XX
PI Schwenk F;
XX
DR WPI; 2001-441873/47.
XX
PT Using site-specific DNA recombinase domain/protein transduction domain
PT fusion proteins for inducing target gene alterations in organisms or
PT cell cultures -
XX
PS Example; Page 63-65; 85pp; English.
XX
CC The present invention relates to use of fusion proteins comprising
CC a site-specific DNA recombinase domain e.g. Cre and a protein
CC transduction domain (PTD) e.g. the Human immunodeficiency virus
CC (HIV) derived TAT peptide, for preparing an agent for inducing
CC target gene alterations in a living organism or cell culture. The
CC present invention also provides a method for inducing gene
CC alterations in living organisms using the fusion proteins of the
CC invention. The present sequence is a pCMV-I-beta-pA vector
CC DNA related to the invention.

SQ Sequence 7332 BP; 1718 A; 1895 C; 1964 G; 1755 T; 0 other;
Query Match 45.7%; Score 2637; DB 22; Length 7332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3131 cgaattcactggccgctcgtttttacaacgctgtagctgggaaaaaccctggcttaccacaac 3190
|||||
Db 4661 cgaattcactggccgctcgtttttacaacgctgtagctgggaaaaaccctggcttaccacaac 4720
|||||
Qy 3191 ttaatcgcccttgacgacacatccccctttcgccagctggcgtaataagcgaagaggcccgca 3250
|||||
Db 4721 ttaatcgcccttgacgacacatccccctttcgccagctggcgtaataagcgaagaggcccgca 4780
|||||
Qy 3251 ccgatacgcccttcccaacagttgcgagcctgaatggcgcaatggcgctgacggtatt 3310
|||||
Db 4781 ccgatacgcccttcccaacagttgcgagcctgaatggcgcaatggcgctgacggtatt 4840
|||||
Qy 3311 ttctccttaagcactctgtgcggtattttcacaccgcataatggtgcactctcagtaacaatct 3370
|||||
Db 4841 ttctccttaagcactctgtgcggtattttcacaccgcataatggtgcactctcagtaacaatct 4900
|||||
Qy 3371 gctctgatgcgcgcagatgtaagccagcccgacacccgcgaacccgctgacgcgccct 3430
|||||
Db 4901 gctctgatgcgcgcagatgtaagccagcccgacacccgcgaacccgctgacgcgccct 4960
|||||
Qy 3431 gacgggctgtctgctcccgcatcccgcttacagacaaagctgtgaccgctctccgggagct 3490
|||||
Db 4961 gacgggctgtctgctcccgcatcccgcttacagacaaagctgtgaccgctctccgggagct 5020
|||||
Qy 3491 gcatgtgtcagaggttttcaccgctcatcaccgaaaacgcgcgagacgaaaggcctcgtga 3550
|||||
Db 5021 gcatgtgtcagaggttttcaccgctcatcaccgaaaacgcgcgagacgaaaggcctcgtga 5080
|||||
Qy 3551 tacgcctattttataggttaatgctatgataataatggtttcttagacgtcaggtggca 3610
|||||
Db 5081 tacgcctattttataggttaatgctatgataataatggtttcttagacgtcaggtggca 5140
|||||
Qy 3611 cttttcggggaaaatgtgcgcgggaacccctattttgtttatttttctaaatacattcaata 3670
|||||
Db 5141 cttttcggggaaaatgtgcgcgggaacccctattttgtttatttttctaaatacattcaata 5200
|||||
Qy 3671 tgtatccgctcatgagacaataaaccttgataaaatgcttcaataatattgaaaaagggaaga 3730
|||||
Db 5201 tgtatccgctcatgagacaataaaccttgataaaatgcttcaataatattgaaaaagggaaga 5260
|||||
Qy 3731 gtatgagtattcaacatttccggtgctgcgcccttattcccttttttgcggtatttgccttc 3790
|||||
Db 5261 gtatgagtattcaacatttccggtgctgcgcccttattcccttttttgcggtatttgccttc 5320
|||||
Qy 3791 ctggttttgcctcaccacagaaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 3850
|||||
Db 5321 ctggttttgcctcaccacagaaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 5380
|||||
Qy 3851 cagcagtggtttacatcgaaactggatctcaacagcgggtaagatccttgagagtttccgcc 3910
|||||
Db 5381 cagcagtggtttacatcgaaactggatctcaacagcgggtaagatccttgagagtttccgcc 5440
|||||
Qy 3911 ccgaagaacggtttcccaatgatgagcacttttaaagtctgctatgtggcggtattat 3970
|||||
Db 5441 ccgaagaacggtttcccaatgatgagcacttttaaagtctgctatgtggcggtattat 5500
|||||
Qy 3971 cccgtattgacgcgcgggcaagagcaactcggctgcgcgcatacactattctcagaatgact 4030
|||||
Db 5501 cccgtattgacgcgcgggcaagagcaactcggctgcgcgcatacactattctcagaatgact 5560
|||||
Qy 4031 tgggtgagtactcaccagtcacagaaaaagcatcttaccggtggcagcagtaagagaat 4090
|||||
Db 5561 tgggtgagtactcaccagtcacagaaaaagcatcttaccggtggcagcagtaagagaat 5620
|||||
Qy 4091 tatgcagtgcctgccataaaccatgagtataacacactgcggcccaacttacttctgacaacga 4150
|||||

Db	5621	tatgcagtgctgccataaaccatgagtgataaacactgcggccaacttacttcttgacaacga	5680
QY	4151	tcggaggaccgaaggagctaacccgcttttttgcacaacaatgggggatcatgttaactcgc	4210
Db	5681	tcggaggaccgaaggagctaacccgcttttttgcacaacaatgggggatcatgttaactcgc	5740
QY	4211	ttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacga	4270
Db	5741	ttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacga	5800
QY	4271	tgctgtagcaatggcaacaacgttgcgcgaactattaaactggcgaaactacttactctag	4330
Db	5801	tgctgtagcaatggcaacaacgttgcgcgaactattaaactggcgaaactacttactctag	5860
QY	4331	cttcccgcccaacaattaatagactggatggagggcggataaaagttagcaggaccacttctgc	4390
Db	5861	cttcccgcccaacaattaatagactggatggagggcggataaaagttagcaggaccacttctgc	5920
QY	4391	gctcggcccttcggcgctggctggtttatttgctgataaaatctggagccggtgagcgtg	4450
Db	5921	gctcggcccttcggcgctggctggtttatttgctgataaaatctggagccggtgagcgtg	5980
QY	4451	ctcgcggtatcatctgcagcactggggccagatcggtaaagccctcccgtatcgtagttatct	4510
Db	5981	ctcgcggtatcatctgcagcactggggccagatcggtaaagccctcccgtatcgtagttatct	6040
QY	4511	acacgacggggagtcaggcaactatggatgaacgaaaatagacagatcgcctgagataggtg	4570
Db	6041	acacgacggggagtcaggcaactatggatgaacgaaaatagacagatcgcctgagataggtg	6100
QY	4571	cctcactgattaaagcattggtaactgtcagaccaagtcttactcataataactttagattg	4630
Db	6101	cctcactgattaaagcattggtaactgtcagaccaagtcttactcataataactttagattg	6160
QY	4631	atttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttgcataactctca	4690
Db	6161	atttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttgcataactctca	6220
QY	4691	tgaccaaaatcccttaacgtgagtttctgcctcactgagcgtcagaccccgtagaaaaa	4750
Db	6221	tgaccaaaatcccttaacgtgagtttctgcctcactgagcgtcagaccccgtagaaaaa	6280
QY	4751	tcaaaaggatcttcttgagatcccttttttctgcgcgtaactctgctgtctgcacaaacaaa	4810
Db	6281	tcaaaaggatcttcttgagatcccttttttctgcgcgtaactctgctgtctgcacaaacaaa	6340
QY	4811	aaccacccgtaccagcgggtggttttgcgcggatcaagagctaccaactcttttcccg	4870
Db	6341	aaccacccgtaccagcgggtggttttgcgcggatcaagagctaccaactcttttcccg	6400
QY	4871	aggtaactggcttcagcagagcgcagataacaaatactgtccttctagttagccgtagt	4930
Db	6401	aggtaactggcttcagcagagcgcagataacaaatactgtccttctagttagccgtagt	6460
QY	4931	taggccaccacttcaagaactctgtagcacgcgcctacataacctcgtctgtctaactcctgt	4990
Db	6461	taggccaccacttcaagaactctgtagcacgcgcctacataacctcgtctgtctaactcctgt	6520
QY	4991	taccagtggtcgtgcagtgccagtgccagtaagtgcgttcttaccggggtggacctcaagacgat	5050
Db	6521	taccagtggtcgtgcagtgccagtgccagtaagtgcgttcttaccggggtggacctcaagacgat	6580
QY	5051	agttaccgggataaggcgagcgggtcgggtgaacgggggttctcgtgcacacagcccagct	5110
Db	6581	agttaccgggataaggcgagcgggtcgggtgaacgggggttctcgtgcacacagcccagct	6640
QY	5111	tgagcgaaacgacctacaccgaaactgagataacctacagcgtgagctatgagaaagcgcca	5170
Db	6641	tgagcgaaacgacctacaccgaaactgagataacctacagcgtgagctatgagaaagcgcca	6700
QY	5171	cgcttcccgaaggagaaaggcgacaggtatccgggtaaagcggcgaggggtcggaacagggag	5230
Db	6701	cgcttcccgaaggagaaaggcgacaggtatccgggtaaagcggcgaggggtcggaacagggag	6760

RESULT 5
AAD04946/C

ID AAD04946 standard; DNA; 8062 BP.

XX AAD04946; AC

XX
DT 17-JUL-2001 (first entry)

XX DE Plasmid pRK76 used to test Cre recombinase mediated inversion.

xx
kw Gene trapping construct; conditional mutation; unidirectional inversion;
kw recombinase recognition sequence; RRS; disruption cassette;
kw selection cassette; transgenic organism; pRK76; Cre recombinase; ds.

XX	
OS	Chimeric - Adenovirus.
OS	Chimeric - ECMV virus.
OS	Chimeric - Unidentified.

XX
PN
WO200129208-A1.

XX
PD 26-APR-2001.

XX
PF 16-OCT-2000; 2000WO-EP10162.

XX
PR 16-OCT-1999; 99EP-0120592.

PR 27-OCT-1999; 99US-0162016.
XX

PA (ARTE-) ARTEMIS PHARM GMBH.
PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.

XX
PI Kuehn R, Von Melchner H, Altschmied J;

XX
DR WPI; 2001-308486/32.

XX
PT
New gene trapping construct capable of causing conditional mutations in

PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped -
XX
PS Example 3; Page 70-73; 78pp; English.
XX
CC The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC recombinase which is capable of unidirectional inversion of double
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is pRK76 vector, which is used to test Cre
CC recombinase mediated inversion, which is related to the invention.
XX
SQ Sequence 8062 BP; 1960 A; 2036 C; 2121 G; 1945 T; 0 other;

Query Match: 45.7%; Score 2637; DB 22; Length 8062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3791 ctgttttctcaccagaaacgctggtgaaagtataaagatgctgaagatcagttgggtg 3850
|||
Dd 7385 CTGTTTTCCTCACCAGAAACGCTGGTGAAGTAAAAAGATGCTGAAGATCAGTTGGGTG 7326

QY 3851 cacgagtggtttacatcgaaactggatctcaacagcggtaagatcccttgagagttttcgcc 3910
|||
Dd 7325 CACGAGTGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTCGCC 7266

QY 3911 ccgaagaacgttttccaatgatgagcacacttttaaaagtctctgatatgtgagcggtattat 3970
|||
Dd 7265 CCGAAGAACGTTTCCAAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGGGTATTAT 7206

QY 3971 cccgtattgacgcggggcaagagcaactcggtcgcccgcatacacactatttcagaatgact 4030
|||
Dd 7205 CCCGTATTGACGCGGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAATGACT 7146

QY 4031 tggttgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaat 4090
|||
Dd 7145 TGGTTGAGTACTACCAGTCAAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT 7086

QY 4091 tatgcagtctgccataaaccatgagtataaactgcggccaacttacttctgacaacga 4150
|||
Dd 7085 TATGCAGTCTGCCATAACCATGAGTGATGAACACTGCGGCCAACTTACTTCTGACAACGA 7026

QY 4151 tcggaggacccgaaggagctaaacgcgttttttgcaacaacatgggggatcatgtaaactcgcc 4210
|||
Dd 7025 TCGGAGGACCGGAAGAGCTAACCCGCTTTTGTGCAACAACATGGGGGATCATGTAACTCGCC 6966

QY 4211 ttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacga 4270
|||
Dd 6965 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGA 6906

QY 4271 tgcctgtagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctag 4330
|||
Dd 6905 TGCCTGTAGCAATGGCAACAACAGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 6846

QY 4331 cttcccgcccaacaattaatagactggatggaggcgggataaaagtgcaggaccacttctgc 4390
|||
Dd 6845 CTTCGCGGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACCTTCTGC 6786

QY 4391 gctcggcccttcggcgtggtgtttattgtctgataaaactctggagccggtgagcgtgggt 4450
|||
Dd 6785 GCTCGGCCCTTCGGGCTGGTGGTTTATTGCTGATAAATCTGGAGCCGCTGAGCGTGGGT 6726

QY 4451 ctcgcggtatcatgtcagcactggggccagatggtaagccctcccgatcgtagttatct 4510
|||
Dd 6725 CTCGCGGTATCATTCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6666

QY 4511 acacgacggggagtcaggcaactatggatgaacgaaaaatagacagatcgtgagataggtg 4570
|||
Dd 6665 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAAAATAGACAGATCGCTGAGATAGGTG 6606

QY 4571 cctcactgattaagcattggttaactgtcagaccaaagtttactcataataacttttagattg 4630
|||
Dd 6605 CCTCACTGATTAAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTG 6546

QY 4631 atttaaaaacttcatttttaatttaaaaggatctaggagaagatcccttttgataaatctca 4690
|||
Dd 6545 ATTTAAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCA 6486

QY 4691 tgaccaaaaatcccttaacgtgagttttctccactgagcgtcagacccccgtagaaaaaga 4750
|||
Dd 6485 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCACCTGAGCGCTCAGACCCCGTAGAAAAAGA 6426

QY 4751 tcaaaaggatctcttgagatccctttttttcttgcgcgtaaatctgctgcttgcaaaacaaaa 4810
|||
Dd 6425 TCAAAAGGATCTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTGCACCAACAAAA 6366

QY 4811 aaccacccgtaccagcgggtgtttgtttgcccggatcaagagctaccactctttttccga 4870
|||
Dd 6365 AACCAACCGCTACCAGCGGTGGTTTGTGTTGCGGATCAAGAGCTACCAACTCTTTTCCGA 6306

QY 4871 aggtaactggccttcagcagagcgcagacataccaaaactgtccttcttagtgcgtagt 4930

3131 cgaattcactgcccgtcgcttttacaaactgctgactgggaaaacccctggcttaccacaac 3190
|||
Dd 8045 CGAATTCACCTGCGCGTCGTTTACAACTGCTGACTGGGAAAACCCCTGGCTTACCCAAC 7986

QY 3191 ttaatgccttcagacacatcccccttttcgacagctgcgtaaatagcgaagagcccgca 3250
|||
Dd 7985 TTAATCGCCTTCAGACACATCCCCCTTTCCGACGCTGGCGTAATAGCGAAGAGGCCCGCA 7926

QY 3251 ccgatacgcccttcccaacagttgcgacgctgaatggcgaatggcgccctgatgcggtatt 3310
|||
Dd 7925 CCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGCAATGGCGCTGATGCGGTATT 7866

QY 3311 ttctccttacgcatctgtcggtattttcacaccgcgcatatggtgcactcagtcacaatct 3370
|||
Dd 7865 TTCTCCTTACGCATCTGTGCGGTATTTTCACCCGCATATGTTGTCACCTCTCAGTACAATCT 7806

QY 3371 gctctgatgcgcgcatagtttaagccagcccgacacccgccaaaccccgctgacgcgccct 3430
|||
Dd 7805 GCTCTGATGCCGATAGTTAAGCCAGCCCGACACCCGCCAACACCCCGCTGACGCGCCT 7746

QY 3431 gacgggcttgctgctcccgccatccgctctacagacaagctgtgacgctctccggagct 3490
|||
Dd 7745 GACGGGCTTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 7686

QY 3491 gcatgtgtcagagggttttccaccgctcatcaccgaaacgcgcgagacgaaaggccctcgtga 3550
|||
Dd 7685 GCATGTGTCAGAGGTTTTCACCGTCAATCCGGAACGCGGAGACGAAAGGGCCCTCGTGA 7626

QY 3551 tacgcctattttataggttaatgtcatgataataatggtttcttagacgtcaggtggca 3610
|||
Dd 7625 TACGCCTATTTTATAGGTTAATGTCAATGATAAATGTTTCTTAGACGTCAGGTGGCA 7566

QY 3611 cttttcggggaaatgtgcgggaacccctatttgtttatttttctaaatacatattcaata 3670
|||
Dd 7565 CTTTTCGGGGAAATGTGCGGGAACCCCTATTGTTTATTTCATAATACATTCAAATA 7506

QY 3671 tgtatccgctcatgagacaataaacctgataaaatgcttcaataataattgaaaaaggaaga 3730
|||
Dd 7505 TGTATCCGCTCATGAGACAATNACCCTGATAAATGCTTCAATAATATATGAAAAGGAAGA 7446

QY 3731 gtatgagtattcaacatttcggtgctgcccttattcccttttttggggcattttgcttc 3790
|||
Dd 7445 GTATGAGTATTCAACATTTCCGCTGCGCCCTTATTCCTCTTTTTCGCGCATTTTGCCTTC 7386

Db 7836 GACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAAGCTGTGACCCGTTCTCCGGGAGCT 7777

QY 3491 gcatgtgtcagagggttttcaccgtcatcacccgaaacgcgcgagacgaaaggccctcgtga 3550

Db 7776 GCATGTGTCAGAGGTTTTCACCGTTCATCCGGAACCGCGGAGACGAAAGGCCTCGTGA 7717

QY 3551 tacgcctattttatagggttaatgtcatgataaataatgggttctttagacgtcaggtggca 3610

Db 7716 TACGCCATATTTTATAGGTTAATGTCTATGATAAATAATGGTTTCTTAGACGTCAGGTGGCA 7657

QY 3611 ctttccggggaaatgtcgcggaaacccctatttggttatttttctaaatacatctcaaata 3670

Db 7656 CTTTTCGGGGAATGTGCGGGAACCCCTATTGTATTATTCTAAATACATTCAAATA 7597

QY 3671 tgtatccgcctcatgagacaataaacctgataaaatgcttcaataatatgaaaaaggaaga 3730

Db 7596 TGTATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATATTGAAAAAGGAAGA 7537

QY 3731 gtagagattatcaacatttcccgtagcgccttatcccttttttgcggcaattttgccttc 3790

Db 7536 GTATGAGTATTCAACATTTCCGTGTGCGCCCTATTCCCTTTTGTGGGCATTTTGCCCTTC 7477

QY 3791 ctggtttttgctcacccagaaaacgctggtgaaagtataaagatgctgaaagatcaagttgggtg 3850

Db 7476 CTGTTTTTGTCAACCCAGAAACGCTGGTGAAGATAAAGATGCTGAAGATCAGTTGGGTG 7417

QY 3851 cacgagtggtttacatcgaactggaatctcaacagcggtaagatcccttgagagtttccgcc 3910

Db 7416 CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC 7357

QY 3911 ccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattat 3970

Db 7356 CCGAAGAACGTTTTCGAATGATGAGCACATTTTAAAGTTCTGCTATGTGCGCGGTATTAT 7297

QY 3971 cccgtattgacgcggggcaagagcaactcggtcgcgcgcatacacactattctcagaatgact 4030

Db 7296 CCCGTATTGACGCGGGCAAGAGCAACTCGTTCGCGCATACACTATTCTCAGAAATGACT 7237

QY 4031 tggttgagtactcaccagtcacagaaaaagcatcttacggatggcgatgacagtaagagaat 4090

Db 7236 TGGTTGAGTACTCACCAAGTACACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA 7177

QY 4091 tatgcagtgctgccataaaccatgaatgataaacactgcggccaaacttacttctgcacaacga 4150

Db 7176 TATGCAGTGTGCCATAAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAACGA 7117

QY 4151 tcggaggaccgaaggagctaaaccgctttttgcacaacatgggggatacatgtaactcgcc 4210

Db 7116 TCGGAGGACCGAAGGAGCTAACCCGCTTTTGTGCAACAACATGGGGGATCATGTAACCTCGCC 7057

QY 4211 ttgatcgttgggaaccggagctgaatgaagcccataccaaacgacgagcgtgacaccacga 4270

Db 7056 TTGATCGTTGGAAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACGA 6997

QY 4271 tgcctgtagcaatggcaacaacgtttgcgaactattaaactggcgaactacttactctag 4330

Db 6996 TGCCGTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 6937

QY 4331 cttcccgcccttcgcggtggttattgctgataaaatctggagccggtgacggaccacttctgc 4390

Db 6936 CTTCCCGGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCCTTCTGC 6877

QY 4391 gctcggcccttcgcggtggttattgctgataaaatctggagccggtgagcgtgggt 4450

Db 6876 GCTCGGCCCTTCCGGCTGGCTGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT 6817

QY 4451 ctcgcgggtatcaattgcagcactggggccagatggttaagccctcccgctatcgtagttatct 4510

Db 6816 CTCGGGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6757

QY 4511 acacgacgggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg 4570

Db 6756 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 6697

QY 4571 cctcactgattaaagcattggtaactgtcagaccaagtttactcatatatacttttagattg 4630

Db 6696 CCTCACTGATTAAAGCATTGGTAACGTGTAGACCAAGTTTACTCATATATACTTTAGATTG 6637

QY 4631 atttaaaacttcatttttaatttaaaaggatctagggtgaagatcccttttttgataaatctca 4690

Db 6636 ATTTAAAACTTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCA 6577

QY 4691 tgaccaaaaatcccttaacgtgagtttcttcgttcacactgagcgtcagaccccgtagaaaaa 4750

Db 6576 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGA 6517

QY 4751 tcaaggatcttcttgagatccctttttctgcgcgtaaatctgctgcttgcaaaacaaaa 4810

Db 6516 TCAAAGGATCTTCTTGAGATCCCTTTTCTGCGGTAATCTGCTGCTTGCAAAACAAAAA 6457

QY 4811 aaccaccgctaccagcggtggttgtttgcggatcaagagctaccaactctttttccga 4870

Db 6456 AACCAACCGCTACCAGCGGTGGTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGA 6397

QY 4871 agttaactggcttcagcagagcgcagatatccaaaatactgtccttctagtagccgtagt 4930

Db 6396 AGTAACTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTCTAGTAGCCGTAGT 6337

QY 4931 taggccaccacttcaagaactctgtagcacccgctacataacctcgtctctgctaactcctgt 4990

Db 6336 TAGGCCACCACCTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCCTGT 6277

QY 4991 taccagtggtctgcccagtggcgataaagtctgtcttaccgggttggactcaagacgat 5050

Db 6276 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTGGACTCAAGACGAT 6217

QY 5051 agttaccggataaaggcgcagcggctcgggtgaacgggggttctcgtgcacacgcccagct 5110

Db 6216 AGTTACCGGATAAGCGCGCAGCGGTTCGGGCTGAACGGGGGGTTCTGTCACACAGCCCCAGCT 6157

QY 5111 tggagcgaacgacctacaccgaactgagataacctacagcgtgagctaagagaaagcgcca 5170

Db 6156 TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAAGCGCCA 6097

QY 5171 cgcttcccgaaagggagaaaaaggcggacaggtatccggtaagcggcagggtcgggaacaggag 5230

Db 6096 CGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGTCGGAACACAGGAG 6037

QY 5231 agcgcacgagggagcttccaggggggaaacgcctgggtatctttatagtcctcgtcggggttc 5290

Db 6036 AGCGCACGAGGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCTCTGCGGGTTTC 5977

QY 5291 gccacctctgacttgagcgtcgatttttggatgctcgtcagggggggagccctatgga 5350

Db 5976 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGCGAGCCCTATGGA 5917

QY 5351 aaaacgccagcaacgcggcctttttacgggttccctggccttttctggccttttgctcacaca 5410

Db 5916 AAAACGCCAGCAACGCGGCCCTTTTACGGTTCTCTGGCCTTTTGTGGCCTTTTGTCTCACA 5857

QY 5411 tgttcttctcgtggttatccctgattctctgggataaccgtattaccgcctttgagtgag 5470

Db 5856 TGTTCTTCTCGGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG 5797

QY 5471 ctgataccgctcgcgcagccgaacgacccgagcgcagcagtcagtcagtcagtcagtcagtcag 5530

Db 5796 CTGATACCGCTCGCCGACGCCGAACGACCGGACGCGCAGTCAGTCAGTCAGTCAGTCAGTCAG 5737

QY 5531 aagagcgcccaatacagcaaacccgctctccccgcgcgttggccgattcattaatgcagct 5590

Db 5736 AAGAGCGCCCAATACGCAAAACCCCTCTCTCCCCGCGCTTGGCCGATTTCATTATGCAGCT 5677

QY 5591 ggcacgacaggtttccccgactggaaaagcgggagtcgagtcgcaacgcaattaatgtgagtt 5650

Db 5676 GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 5617

Qy	5651	agctcactcattagggcaccgccagcgtttacac	tttatgcttccggctcgatg	tggtg	5710
Db	5616	AGCTCACTCATAGGCACCCAGGCTTTACACTT	TATGCTTCCGGCTCGTATGTTGTG		5557
Qy	5711	gaattgtgagcggataacaatttcacacaggaac	agctatgaccatgattacgcca		5767
Db	5556	GAATTGTGAGCGGATAACAAATTTACACAGGAAC	AGCTATGACCATGATTACGCCA		5500

RESULT 7
AAZ20086
ID AAZ20086 standard; DNA; 8854 BP.
XX
XX AC AAZ20086;
XX
XX
DT 05-JAN-2000 (first entry)
XX
XX
DE DNA encoding chimeric cry1B insecticide HyFLIB.
XX
XX
KW Insecticide; HyFLIB; crystal protein; delta-endotoxin; toxin;
KW cry1B; cryIA(b); maize; transgenic plant; European corn borer;
KW Ostrinia nubilalis; entomocide; crop protection; biological control;
KW PCIB5520; ss.

PN WO9950293-A1.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-EP02175.
 XX
 PR 01-APR-1998; 98US-0053549.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 XX
 PI Desal NM;
 XX
 DR WPI; 1999-601323/51.
 DR P-PSDB; AAY31990.
 XX
 PT Chimeric insecticidal protein comprising cryIB and cryIA(b) portions
 XX Claim 3; Page 34-42; 85pp; English.
 PS

```
Query Match      45.7%; Score 2637; DB 20; Length 8854;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	3131	cgaattcactggccgtcgcttttacaacgtcgtgactgggaaaaacccctggcggttacccaac	3190
Db	3892	cgaattcactggccgtcgcttttacaacgtcgtgactgggaaaaacccctggcggttacccaac	3951
QY	3191	ttaatcgcccttgccagcacatcccccctttcgccagctggcgttaatagcgaagagggccgcga	3250
Db	3952	ttaatcgcccttgccagcacatcccccctttcgccagctggcgttaatagcgaagagggccgcga	4011
QY	3251	ccgatacgcccttcccaaacagttgcgccagccctgaatggcggaatggcgccctgatgcggtatt	3310
Db	4012	ccgatacgcccttcccaaacagttgcgccagccctgaatggcggaatggcgccctgatgcggtatt	4071
QY	3311	ttctcccttacgcatactgtgcggtattttcacaccgcataatggtgcactctcagttacaattct	3370
Db	4072	ttctcccttacgcatactgtgcggtattttcacaccgcataatggtgcactctcagttacaattct	4131
QY	3371	gctctgatccgcgcatagtttaagccagcccccgacacccgccaaaccccgctgacgcgcacct	3430
Db	4132	gctctgatccgcgcatagtttaagccagcccccgacacccgccaaaccccgctgacgcgcacct	4191
QY	3431	gacgggcttgctgctccccgcatccgccttacagacaagctgtgacccgtctccgggagct	3490
Db	4192	gacgggcttgctgctccccgcatccgccttacagacaagctgtgacccgtctccgggagct	4251
QY	3491	gcatagtgtcagaggttttccacgcgtcatcacccgaaacgcgcgagacgaaagggcctcgtga	3550
Db	4252	gcatagtgtcagaggttttccacgcgtcatcacccgaaacgcgcgagacgaaagggcctcgtga	4311
QY	3551	tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggca	3610
Db	4312	tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggca	4371
QY	3611	cttttcggggaaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaataa	3670
Db	4372	cttttcggggaaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaataa	4431
QY	3671	tgtatccgcctcatgagacaataaacctgataaaatgcttcaataatattgaaaaagggaaga	3730
Db	4432	tgtatccgcctcatgagacaataaacctgataaaatgcttcaataatattgaaaaagggaaga	4491
QY	3731	gtatgagtattcaacatttcgcgtgcgcccttattcccttttttgcggcattttgccttc	3790
Db	4492	gtatgagtattcaacatttcgcgtgcgcccttattcccttttttgcggcattttgccttc	4551
QY	3791	ctgtttttgctcaccagaaaacgtggtgaaagttaaagatgctgaagatcagttgggtg	3850
Db	4552	ctgtttttgctcaccagaaaacgtggtgaaagttaaagatgctgaagatcagttgggtg	4611
QY	3851	cacgagtgggttacatcgaactggatctcaacagcggtaagatcccttgagagttttcgcc	3910
Db	4612	cacgagtgggttacatcgaactggatctcaacagcggtaagatcccttgagagttttcgcc	4671
QY	3911	ccgaagaacgttttcccaatgataagcacactttttaaagttctgctatgtggcgcggtattat	3970
Db	4672	ccgaagaacgttttcccaatgataagcacactttttaaagttctgctatgtggcgcggtattat	4731
QY	3971	cccgtattgacgcggggcaagagcaactcggtcgccgcatacactattctcagaatgact	4030
Db	4732	cccgtattgacgcggggcaagagcaactcggtcgccgcatacactattctcagaatgact	4791
QY	4031	tggttgagtactaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat	4090
Db	4792	tggttgagtactaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat	4851
QY	4091	tatgcagtgtgccataaccatgagtataaacactgcggccaacttacttcttgacaacaga	4150
Db	4852	tatgcagtgtgccataaccatgagtataaacactgcggccaacttacttcttgacaacaga	4911
QY	4151	tcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatactgttaactcgcc	4210
Db	4912	tcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatactgttaactcgcc	4971

QY 4758 atctcttgagatccttttttctgcgctaatactgctgcttgcacacacacacacac 4817
|||||
Db 4675 ATCTTCTTGAGATCCTTTTCTGCGCTAATCTGCTGCTTGCACACAAAAACCA 4616
XX
QY 4818 gctaccagcgggtggtttgttgcggatcaagagctaccacactcttttccgaaggtaac 4877
|||||
Db 4615 GCTACCAGCGGTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC 4556
QY 4878 tggcttcagcagcagcagcagataccacataactgtccttctagtgtagcgttagtgcca 4937
|||||
Db 4555 TGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCGTAGTTAGGCCA 4496
QY 4938 ccacttcaagaactctgtagcacccgctacatacctcgtctgctaactcctgttaccagt 4997
|||||
Db 4495 CCACCTTCAAGAACTCTGTAGCACCGCTACATACTCGCTCTGCTAATCTGTTACCAGT 4436
QY 4998 ggctgctgcagtgccgataagtctgtcttaccgggttgactcaagacgatatgtacc 5057
|||||
Db 4435 GGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTGGACTCAAGACGATAGTTACC 4376
QY 5058 ggataaggcgcagcggctcgggctgaacgggggttcgtgcacacagccccagcttggagcg 5117
|||||
Db 4375 GGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTCTGTCACACACAGCCACGCTTGGAGCG 4316
QY 5118 aacgacctacacccgaactgagatacctacagcgtgagctatgagaaagcgcacgcttcc 5177
|||||
Db 4315 AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCC 4256
QY 5178 cgaagggaagagcggacaggtatccggtaagcggcagggcgagcctatggaacagcgcac 5237
|||||
Db 4255 CGAAGGGAGAAAGCGCGACAGGTATCCGGTAAGCGGAGGTCGGAACAGGAGAGCGCAC 4196
QY 5238 gagggagcttcaggggggaaacgcctggtatcttatagtcctgctgggttccgccacct 5297
|||||
Db 4195 GAGGGAGCTTCCAGGGGGAAACGCCCTGCTATCTTTATAGTCTCTGCGGTTTCGCCACCT 4136
QY 5298 ctgacttgagcgtcgatttttctgctgctcgtcagggggcgagcctatggaacacgc 5357
|||||
Db 4135 CTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGGCTATGGAACACGC 4076
QY 5358 cagcaacggcgccctttttacggcttcctggccttttgcggccttttgcctcacatgttctt 5417
|||||
Db 4075 CAGCAACGGCGCCTTTTACGGTTCTGCGCCTTTTGTGGCCTTTTGTCTCACATGTTCTT 4016
QY 5418 tctgcttatccccctgattctgtggaataaccgtattaccgccttgcagtgagctgatac 5477
|||||
Db 4015 TCCTGCGTTATCCCTCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC 3956
QY 5478 cgctgcgcagcagccgaacgacgagcagcagcagtcagtcagtcagtcagtcagtcagtcag 5537
|||||
Db 3955 CGCTCGCGCAGCGCGGAACGACCGAGCGGAGTCAGTGAGCGAGGAGCGGAAGAGCG 3896
QY 5538 cccaatacgaacacccgcctctcccccgcgcttggccgattcattaatgcagctggcacga 5597
|||||
Db 3895 CCCAATACGCAAAACCGCCTCTCCCCCGCGGTGGCCGATTTCATTATGCAGCTGGCACGA 3836
QY 5598 caggtttccgactggaaagcggcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 5657
|||||
Db 3835 CAGGTTTCCGACTGGAAAGCGGCGAGTCAGCGCAACGCAATTAATGTAGTTAGCTCAC 3776
QY 5658 tcattaggcaccacccagcctttacactttatgcttccggctcgtatgtgtgtggaattgt 5717
|||||
Db 3775 TCATTAGGCACCCACAGGCTTTACACTTTATGCTTCCGGCTCGTATGTGTGTGGAATTGT 3716
QY 5718 gagcggataacaatttcacacaggaacagctatgaccatgattacgcca 5767
|||||
Db 3715 GAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGATTACGCCA 3666

RESULT 9
AAD10238/c

ID AAD10238 standard; DNA; 7164 BP.

XX

AC AAD10238;
XX
DT 24-SEP-2001 (first entry)
XX
DE Commercial plasmid vector pCMVbeta.
XX
KW Plasmid; cyclic; circular; replicon; exogenous gene; marker gene;
transcription termination; immunostimulatory sequence; ISS; antiviral;
non-essential nucleotide; molecular biology application; gene therapy;
DNA vaccine; cloning; gene expression; in vitro protein production;
cytostatic; pCMVbeta; cytomegalovirus promoter; lacZ gene;
ampicillin gene; simian virus 40; SV40 intron; ds.
XX
OS Chimeric - Cytomegalovirus.
OS Chimeric - Rhesus macaque polyoma virus
OS Chimeric - Unidentified.
FH Key Location/Qualifiers
FT misc_feature 1..26
FT /*tag= a
FT /note= "Corresponds to the non-essential nucleotide
sequence that have been removed in the novel DNA
plasmid vector"
FT 4513..4890
FT /*tag= b
FT /note= "pUC18 lacZ promoter that corresponds to the
non-essential nucleotide sequence that have been removed
in the novel DNA plasmid vector"
FT 7007..7164
FT /*tag= c
FT /note= "5' non-coding region of lacZ gene from pUC18 that
corresponds to the non-essential nucleotide sequence
that have been removed in the novel DNA plasmid vector"
XX
PN WO200151626-A2.
XX
PD 19-JUL-2001.
XX
PF 09-JAN-2001; 2001WO-US01255.
XX
PR 10-JAN-2000; 2000US-0480879.
XX
PA (ELIM-) ELIM BIOPHARMACEUTICALS INC.
PI Lu X, Sun L, Zhang Y;
XX
DR WPI; 2001-451855/48.
XX
PT New plasmid DNA vectors, useful for most molecular biology
applications, e.g. gene therapy, DNA vaccines, cloning and expression
of genes, and in the in vitro production of polypeptides and/or
proteins
XX
PS Example 2; Page 42-44; 50pp; English.
XX
CC The present invention relates to plasmid DNA vectors comprising
essentially of a replicon and at least one other component selected from
promoter, intron, exogenous gene, transcription termination sequence,
selectable marker gene, detectable marker gene and an immunostimulatory
sequence (ISS), where the non-essential nucleotide sequences have been
substantially removed from these vectors. The plasmid DNA vectors are
useful in most molecular biology applications, e.g. gene therapy, DNA
vaccines, cloning and expression of genes, and in the in vitro production
of polypeptides and/or proteins. The present sequence is a commercial
plasmid DNA vector pCMVbeta which comprises Cytomegalovirus promoter,
Simian Virus 40 (SV40) intron, the lacZ gene, pUC origin of replication
and the ampicillin gene.
XX
SQ Sequence 7164 BP; 1700 A; 1840 C; 1897 G; 1727 T; 0 other;

Query Match 45.6%; Score 2630; DB 22; Length 7164;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	3138 actggccgtcggtttacaacgctcgtagctgggaaaaacccctggcggttaccacaacttaatcg 3197
Db	7164 ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACCTTAATCG 7105
QY	3198 ccttgccagcacatccccctttccgcccagctggcgtaataagcgaagagcccgcaccgateg 3257
Db	7104 CCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCG 7045
QY	3258 cccttcccaacagttgcgccagccctgaatggcgcaatggcgccctgatgcggtattttctcct 3317
Db	7044 CCCTTCCCAACAGTTGCCGAGCCTGAATGGCGAATGGCGCTGATGCGGTATTTTCTCCT 6985
QY	3318 tacgcattctgtgcggtatttcacaccgcataatggtgcactctcagtaacaatctgctctga 3377
Db	6984 TACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACTCTCAGTACAATCTGCTCTGA 6925
QY	3378 tgccgcatagttaagccagccccccgcacacccgcacccgcctgacgcgccctgacgggc 3437
Db	6924 TGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGC 6865
QY	3438 ttgtctgctccggcatccgcttacagacaagctgtgacgctctccgggagctgcatgtg 3497
Db	6864 TTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTG 6805
QY	3498 tcagaggttttcacctgcatcaccgaaacgcgcgagacgaagggccctcgtgatacgcct 3557
Db	6804 TCAGAGGTTTTTCAACCGTCATCACCGAAACGCCGGAGACGAAAGGCCCTCGTGATACGCC 6745
QY	3558 atttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcaacttttcg 3617
Db	6744 ATTTTATAGTTAATGTCAATGATAATAATGGTTTCTTAGACGTGAGGTGGCACTTTTCG 6685
QY	3618 gggaaatgtgcgggaacccctatttgtttatttttctaaatacattcaaaatatgtatcc 3677
Db	6684 GGGAAATGTGCGCGGAACCCCTATTGTTTTATTTTCTAAATACATTCAAATATGTATCC 6625
QY	3678 gctcatgagacaaataacccctgataaaatgcttcaataatttgaaaaagggaagagtatgag 3737
Db	6624 GCTCATGAGACAAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAG 6565
QY	3738 tattcaacatttcggtgcgcccttattcccttttttggggcattttgccttccctgttt 3797
Db	6564 TATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTGTGGCGCATTTTGCCCTTCCTGTTTT 6505
QY	3798 tgctcaccagaaacgcgtggtgaaagtaaaagatgctgaagatcagttg9gtgcacgagt 3857
Db	6504 TGCTCACCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGT 6445
QY	3858 g9gtttacatcgaaactggatctcaacacgcggtgaagatccctgagagttttcgccccgaaga 3917
Db	6444 GGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCCTTGAGAGTTTTCGCCCCGAAGA 6385
QY	3918 acgttttccaaatgatgagcaacttttaaagtctgctatgtggcgcggtattatcccgat 3977
Db	6384 ACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTAT 6325
QY	3978 tgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggtga 4037
Db	6324 TGACGCCGGGCAAGAGAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGA 6265
QY	4038 gtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagaaattatgcag 4097
Db	6264 GTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGTCAG 6205
QY	4098 tgctgccataaacatgagtgataaacactgcggcccaacttacttctgacaacgatcggagg 4157
Db	6204 TGCTGCCATAACCATGAGTGATAACACTGCGGCCCACTTACTTCTGACAACGATCGGAGG 6145
QY	4158 accgaaggagctaaccgcttttttgcaacaacatgggggatcatgtaaactcgcttgatcg 4217
Db	6144 ACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCG 6085

QY 5298 ctgacttgagcgtcgatttttgtgatgctcgtcaggggggagcctatggaaaaacgc 5357
|||||
Db 5004 CTGACTTGAGCGTCGATTTTGTGATGCTCCTCAGGGGGGCGGAGCCCTATGGAAAACGC 4945
QY 5358 cagcaaacgcgccttttttacggttccctgtgccccttttgccttttgcctcacatgttctt 5417
|||||
Db 4944 CAGCAACGGCGCTTTTACGGTTCTCTGGCTTTTGTGGCTTTTGTCTCACATGTTCTT 4885
QY 5418 tcctgcgttatccoctgattctgtggataaccgtattaccgcctttgagtgagctgatac 5477
|||||
Db 4884 TCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC 4825
QY 5478 cgctcgcgcagccgaacgacgagcgcagcagtcagtgagcgcgaggaagcgagcg 5537
|||||
Db 4824 CGCTCGCCGCGAGCGGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAACGGAAGCG 4765
QY 5538 cccaatacgcacaaccgcctctccccgcgcttgccgattcattaatgcagctggcacga 5597
|||||
Db 4764 CCCAATACGCAAAACCGCCTCTCCCCGCGCTTGGCCGATTCAATTAATGAGCTGGCACGA 4705
QY 5598 caggtttcccgcactggaaagcgggcagtgagcgcgaacgcaattaatgtgagttagctcac 5657
|||||
Db 4704 CAGGTTTCCCGACTGGAACCGGGCAGTCAGCGCAACGCAATTAATGTGAGTTAGCTCAC 4645
QY 5658 tcattaggcaccgcagcgttttacacctttatgcttccggctcgtatgttgtggaattgt 5717
|||||
Db 4644 TCATTAGGCACCCAGCGCTTTACACTTATGCTTCCGGCTCGTATGTTGTGGAATTGT 4585
QY 5718 gagcggataacaatttcacacaggaacagctatgaccatgattacgcca 5767
|||||
Db 4584 GAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGATTACGCCA 4535

RESULT 10
AAS00153/C
ID AAS00153 standard; DNA; 7383 BP.
XX AAS00153;
AC AAS00153;
XX
DT 24-MAY-2001 (first entry)
XX
DE Matrix metalloprotease 9 (MMP9) promoter:Beta galactosidase construct.
XX
KW Human; matrix metalloprotease 9; MMP9; Beta galactosidase; lacZ;
KW skin metabolism-related promoter; transgenic animal; hair removal;
KW cosmetic; retinoid; transforming growth factor beta; TGF beta;
KW wound healing; transforming growth factor alpha; TGF alpha; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
XX
PN WO200117343-A1.
XX
PD 15-MAR-2001.
XX
PF 08-SEP-2000; 2000WO-US24637.
XX
PR 10-SEP-1999; 99US-0393483.
XX
PA (GENO) GEN HOSPITAL CORP.
XX
PI Burgeson R, Amano S, Kishimoto J, Nishiyama T, Ehama R;
XX
DR WPI; 2001-235146/24.
XX
PT Evaluating a compound for its effect on skin, by administering the
PT compound to a transgenic animal e.g. mice having a reporter gene
PT coupled to skin-metabolism promoter and evaluating expression of the
PT reporter gene -
XX
PS Example 9; Fig 4; 70pp; English.
XX
CC The sequence represents a promoter construct consisting of human

CC matrix metalloprotease 9 (MMP9) promoter fused to Escherichia coli
CC beta galactosidase (lacZ). The construct is used in the evaluation
CC of a treatment for its effect on skin, comprising providing a transgenic
CC animal having a reporter gene encoding a luminescent or fluorescent
CC product, coupled to a skin metabolism-related promoter, administering the
CC treatment to the transgenic animal and evaluating expression of the
CC reporter gene. The method is useful for evaluating a treatment e.g. the
CC removal of hair (e.g. by plucking or shaving), or the administration of a
CC compound, for its effect on skin. The compound is a cosmetic, a non-toxic
CC substance, a substance approved for human drug or cosmetic use in one or
CC more jurisdictions, a retinoid or its derivative, transforming growth
CC factor beta or alpha (TGF beta/alpha). The method is also useful for
CC determining the stage of the hair cycle in an animal which expresses a
CC reporter molecule in hair follicle. The transgenic animals are useful for
CC evaluating the effect of a treatment on wound healing and for analysing
CC the expression of a transgene.
XX
SQ Sequence 7383 BP; 1740 A; 1913 C; 1999 G; 1731 T; 0 other;

Query Match 45.6%; Score 2630; DB 22; Length 7383;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3138 actggccgtcgttttacaaacgtcgtgactggggaacccctggcgttacccaaacttaacg 3197
|||||
Db 7383 ACTGGCCGTCGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATCG 7324
QY 3198 ccttgacgacacatcccccttccgacgtggtggaatggcgctgatgcggtattttctcct 3317
|||||
Db 7263 CCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTGATGCGGTATTTTCTCCT 7204
QY 3318 tacgcatctgtgcgggtatttcacacccgcataatggtgcactctcagtaacaatctgctga 3377
|||||
Db 7203 TACGCATCTGTGCGGTATTTTCACACCCGCATATGTTGTCACCTCAGTACAATCTGCTCTGA 7144
QY 3378 tgccgcatagtttaagccagccccgcagacacccgcgaacaccccgctgacgcgccctgacgggc 3437
|||||
Db 7143 TGCCGCATAGTTAAGCCAGCCCCGACACCCCGCAACACCCCGCTGACGCGCCCTGACGGGC 7084
QY 3438 ttgtctgctcccgcatccgcttacagacaagctgtgaccgtctccggagctgcatgtg 3497
|||||
Db 7083 TTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTG 7024
QY 3498 tcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcct 3557
|||||
Db 7023 TCAGAGGTTTTCACCGTCATCACCCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCT 6964
QY 3558 attttataggttaatgcatgataataatgggtttcttagacgtcaggtggcacttttcg 3617
|||||
Db 6963 ATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTGAGGTGGCAGCTTTTCG 6904
QY 3618 gggaaatgtgcgcggaaacccctatttgtttatttttctaataacattcaaatatgtatcc 3677
|||||
Db 6903 GGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCC 6844
QY 3678 gctcatgagacaataaccctgataaaatgcttcaataataattgaaaaaggaagatgatgag 3737
|||||
Db 6843 GCTCATGAGACAATAACCCCTGTATAAATGCTTCAATAATATTGAAAAAAGGAAGATATGAG 6784
QY 3738 tattcaacatttcctgctgcgcccttattcccttttttgcggcattttgccttcctgtttt 3797
|||||
Db 6783 TATTCAACATTTCCGTGTGCGCCCTTATCCCTTTTTCGGCATTTTTCCTTCTCCTGTTT 6724
QY 3798 tgctcaccagaaacgcgtggtgaaagtaaaagatgctgaagatcagttgggtgcacgagt 3857
|||||
Db 6723 TGCTCACCCACAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT 6664
QY 3858 gggttacatcagtaactggatctcaacagcggtaagatccttgagagtttttcgccccgaaga 3917

Db	3115	AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAAG	3055
Qy	4757	gatcttcttgagatccctttttctgcgcgtaaatctgctgtcttgcaaacacacacacacacac	4816
Db	3055	GATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGTGCAACACAAAAAACCCAC	2996
Qy	4817	cgctaccagcgggtgttgtttgcccggatcaagagctaccacactcttttccgaaggtaa	4876
Db	2995	CGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA	2936
Qy	4877	ctggcttcagcagagcgcagataccaaaatactgtccttctagtgtagcgttagtgcc	4936
Db	2935	CTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCC	2876
Qy	4937	accacttcaagaactctgtagcacccgctacatacctcgtctctgctaactcctgtttaccag	4996
Db	2875	ACCACITCAAGAACTCTGTAGCACCGCTACATACTCGCTCTGCTAATCCTGTATTACCAG	2816
Qy	4997	tggctgctgccagtggcgataaagtctgtcttaccggggttggactcaagacgatatgttac	5056
Db	2815	TGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACC GG GTTGGACTCAAGACGATAGTTAC	2756
Qy	5057	cggataaaggcgcagcggctcgggctgaacgggggggttcgtgcacacagccacgcttggagc	5116
Db	2755	CGGATAAAGCGCAGCGGTTCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGC	2696
Qy	5117	gaacgacctacacccaactgagatacctacacgctgagctatgagaaagcgcacgcttc	5176
Db	2695	GAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC	2636
Qy	5177	ccgaaggagaaaggcggacaggtatcccggtaaagcggcagggtcggaacagggagagcgca	5236
Db	2635	CCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAAACAGGAGAGCGCA	2576
Qy	5237	cgaggggagcttccaggggggaaacgcctggatatctttatagtcctgtcgggttccgccacc	5296
Db	2575	CGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCTCTCGGGTTCGCCACC	2516
Qy	5297	tctgacttgagcgtcgatcttttgtgatgctcgtcagggggcgagccttatggaaaaaacg	5356
Db	2515	TCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACG	2456
Qy	5357	ccagcaacgggccttttacgggttcctggccttttgcggccttttgcctcacatgtct	5416
Db	2455	CCAGCAACGGCGCCTTTTACGGTTCTTGCCCTTTTGTGGCCCTTTTGTCTACATGTTCT	2396
Qy	5417	ttcctgcgttatccctgattctctgtggataaccgtattaccgccttttgagctgagctgata	5476
Db	2395	TTCTCTGCGTTATCCCTCTGATTCCTGCGATACCGGTATTACCGCCTTTGAGTGAGCTGATA	2336
Qy	5477	ccgctcgcgcagccgaacgacgagcgcagcgcagtcagtcagtcagtcagtcagtcagtcagtc	5536
Db	2335	CCGCTCGCGCAGCGGAACGACCGAGCCGACGGAGTCACTGAGCGGAGGAAGAGC	2276
Qy	5537	gccccatacgcaaacccctctccccgcgcttgcccgattcattaatgtgaggttagctca	5596
Db	2275	GCCCAATACGCAAAACCGCCTCTCCCCCGCGGTGGCCCGATTTCATTAAATGCAGCTGGCAGC	2216
Qy	5597	acaggttccgcagctggaagcgggcagtgagcgcgaacgcaattaatgtgaggttagctca	5656
Db	2215	ACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA	2156
Qy	5657	ctcattaggcaccgccaggcttttacactttatgcttccggctcgatgttctgtgtggaattg	5716
Db	2155	CTCATTTAGGCACCCCGAGCTTTACACTTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTG	2096
Qy	5717	tgagcggataacaaatttcacacagggaaaaacgctatgacctatgattacg	5764
Db	2095	TGAGCGGATAACAATTCACACAGGAACAGCTATGACCATGATTACG	2048

AAC82936/c

ID AAC82936 standard; DNA; 4950 BP.

XX

AC AAC82936;

XX

DT 02-APR-2001 (first entry)

XX

DE Transdominant effector peptide screening DNA #8.

XX

KW Intracellular transdominant bioactive agent; screening; cell phenotype;

KW effector peptide; ss.

XX

OS Unidentified.

XX

PN US6153380-A.

XX

PD 28-NOV-2000.

XX

PF 23-JAN-1997; 97US-0789333.

XX

PR 23-JAN-1996; 96US-0589108.

PR 23-JAN-1996; 96US-0589911.

XX

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PA (RIGE-) RIGEL PHARM INC.

XX

PI Rothenberg SM, Nolan GP;

XX

DR WPI; 2001-060084/07.

XX

PT Methods for screening intracellular transdominant effector peptides and

PT RNA molecules comprise delivering random oligonucleotides to cells,

PT which are then screened for an altered phenotype -

XX

PS Example 2; Column 81-88; 57pp; English.

XX

CC This invention describes novel in vitro screening methods (I) for a

CC transdominant intracellular bioactive agent capable of altering the

CC phenotype of a cell. (I) comprises: (a) introducing a molecular library

CC of randomized candidate nucleic acids into several cells; and (b)

CC screening the cells for a cell exhibiting an altered phenotype, where the

CC altered phenotype is due to the presence of a transdominant bioactive

CC agent. The methods are particularly useful for screening intracellular

CC transdominant effector peptides and RNA molecules selected inside living

CC cells from randomized pools. (I) is also useful for introducing random

CC libraries into cells to screen for bioactive compounds. The methods allow

CC rapid and highly efficient screening of large numbers of random

CC oligonucleotides and their corresponding expression products in a single

CC step. In addition, the methods allow screening in the absence of

CC significant prior characterization of the cellular defect.

XX

SO Sequence 4950 BP; 1163 A; 1347 C; 1256 G; 1184 T; 0 other;

QY 3377 atgccgcataagtaagccagccccgacaccccgccaaacaccccgctgacgcgccctgacggg 3436
|||||
Db 4435 ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCGCAACACCCGCTGACGCGCCCTGACGGG 4376

QY 3437 ctgtctgctcccgccatccgcttacagacaagctgtgacccgtctccgggagctgcatgt 3496
|||||
Db 4375 CTGTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGT 4316

QY 3497 gtcagagggttttcacccgctcatcacccgaacgcgcgagacgaaaggccctcgtgatacgc 3556
|||||
Db 4315 GTCAGAGGTTTTCACCGTTCATCACCCGAACGCGCGAGACGAAAGSGCCTCGTGATACGCC 4256

QY 3557 tatttttatagggttaagtcatgataataatgggtttcttagacgtcagggtggcacttttc 3616
|||||
Db 4255 TATTTTATAGGTTAATGTCAATGATATAATAATGGTCTTCTTAGACGTCAGGTGGCACTTTTC 4196

QY 3617 ggggaaatgtgcgcggaaacccctattgtttatttttctaaatacattcaaatatgtatc 3676
|||||
Db 4195 GGGGAAATGTGCGCGGAACCCCTATTGTGTTATTTTCTAAATACATTCAAATATGTATC 4136

QY 3677 cgctcatgagacaataaccctgataaatgcttcaataattgaaaagaaagatatga 3736
|||||
Db 4135 CGCTCATGAGACAATAACCTGTATAAATGCTCAATAATATTGAAAAAGGAAGATGA 4076

QY 3737 gtattcaacatttccggtgtcgccccatttcccttttttgcggcattttgcccctcctgttt 3796
|||||
Db 4075 GTATTCAACATTTCGGTGTGCCCCTTATTCCTTTTTTGGCGCATTTTGCCCTTCCTGTTT 4016

QY 3797 ttgctcaccagaaaacgctgggtgaaagtaaaagatgctgaagatcagttgggtgcacgag 3856
|||||
Db 4015 TTGCTCACCCAGAAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 3956

QY 3857 tgggttacatcgaaactggatctcacaacagcgggtaagatccttgagagtttccgccccgaag 3916
|||||
Db 3955 TGGGTTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCCGAAG 3896

QY 3917 aacgttttccaatgatgagcacttttaagttctgctatgtggcgcggtattatccccgta 3976
|||||
Db 3895 AACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGTATTATCCCCGTA 3836

QY 3977 ttgacgccgggcaagagcaactcgggtcgcgccatacacactattctcagaatgacttggtg 4036
|||||
Db 3835 TTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTG 3776

QY 4037 agtactcaccagtcacagaaaagcatcttacggatggcatgacagtgaagaaattatgca 4096
|||||
Db 3775 AGTACTCACCACTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAAGAGAAATTATGCA 3716

QY 4097 gtgctgccataaaccatgagtataaacactcggcgccaaacttacttcttgacaaacgatcggag 4156
|||||
Db 3715 GTGCTGCCATAACCATGAGTGATAACACTGCGGCCCAACTTACTTCTGACAAACGATCGGAG 3656

QY 4157 gaccgaaggagctaacccgttttttgcaacaacatgggggatcatgtaaactcgccttgatc 4216
|||||
Db 3655 GACCGAAGGAGCTAACCGCTTTTGTGCAACATGGGGGATCATGTAACTCGCCTTGATC 3596

QY 4217 gttgggaacccgagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctg 4276
|||||
Db 3595 GTTGGGAACCCGAGCTGAATGAAGCCATACCAAAACGACGAGCGGTGACACCACGATGCCTG 3536

QY 4277 tagcaatggcaacaacgcttgcgcaaacactattaaactggcgaactacttactctagcttccc 4336
|||||
Db 3535 TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC 3476

QY 4337 ggcaacaattaatagactggatggaggcgagataaagttgcaggaccacttctgcgtcgg 4396
|||||
Db 3475 GGCAACAATTAATAGACTGGATGGAGCGCGGATAAAGTTGCAGGACCACCTTCTGCGCTCGG 3416

QY 4397 cccctccggctgggttttattgctgataaaatctggagccgggtgagcgtgggtctcgcg 4456
|||||
Db 3415 CCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCCGGTGAGCGTGGGTCTCGCG 3356

QY 4457 gtatcattgcagcactggggccagatggtaagccctcccgatatcgtagttatctacacga 4516
|||||
Db 3355 GTATCATTTGCAGCACTGGGGCCAGATGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 3296

QY 4517 cggggagtcaggcaactatggatgaacgaaatagacagatcgcgtgagataggtgcctcac 4576
|||||
Db 3295 CGGGGAGTCAAGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGTGCCTCAC 3236

QY 4577 tgattaagcatttgtaactgtcagaccaaagtttactcatatatacttttagattgatttaa 4636
|||||
Db 3235 TGATTAAGCATTGGTAACCTGTAGACCAAGTTTACTCATATATATACTTTAGATTGATTTAA 3176

QY 4637 aacttcatttttaatttaaaggatcttaggtgaagatccctttttgataatctcatgacca 4696
|||||
Db 3175 AACTTCATTTTAAATTAAAAGGATCTAGTGAAGATCCCTTTTGTGATAATCTCATGACCA 3116

QY 4697 aaatcccttaacgtgagtttctgcctccactgagcgtcagacccccctagaaaaagatcaaaag 4756
|||||
Db 3115 AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAG 3056

QY 4757 gatcttcttgagatcccttttctgcgcgtaaatctgctgcttgcaaaaacaaaaaacccac 4816
|||||
Db 3055 GATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAACCCAC 2996

QY 4817 cgctaccagcgggtggtttgtttgcggatcaaacataactgtccttctagtgtagccgttagggcc 4876
|||||
Db 2995 CGTACCAGCGGTGGTGTGTTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA 2936

QY 4877 ctggcttcagcagagcgcagatataccaaataactgtccttctagtgtagccgttagtgcc 4936
|||||
Db 2935 CTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCCGTAGTTAGGCC 2876

QY 4937 accacttcaagaactctgtagcacccgctacatacctcgtcctgtcctaactcctgttaccag 4996
|||||
Db 2875 ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCTCTTACCAG 2816

QY 4997 tggctgctgccagtggcgataaagtcgtcttaccgggttggactcaagacgatatgttac 5056
|||||
Db 2815 TGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACC GGTTGGACTCAAGACGATAGTTAC 2756

QY 5057 cggataaagcgcagcggctcgggctgaacgggggttcgtgcacacagccccagcttggagc 5116
|||||
Db 2755 CGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCGCTTGGAGC 2696

QY 5117 gaacgacctacacccgaactgagataacctacacgctgagctatgagaaagcgcacgcttc 5176
|||||
Db 2695 GAACGACCTACACCCGAACCTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC 2636

QY 5177 ccgaagggaagaaaggcgacaggtataccggtaagcggcagggctcggaacagagagcgca 5236
|||||
Db 2635 CCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGGGTCGGAACAGGAGAGCGCA 2576

QY 5237 cgagggagcttccagggggaaacgcctgggtatcttatagtcctgtcgggtttcgcacc 5296
|||||
Db 2575 CGAGGGAGCTTCCAGGGGAAACGCCCTGCTATCTTATAGTCCCTGTCGGGTTCGCCACC 2516

QY 5297 tctgacttgagcgtcgatattttgtgatcgtcaggggtcagggggtcctatggaataaacg 5356
|||||
Db 2515 TCTGACTTGAGCGTCGATTTTGTGTGATGCTGCTCAGGGGGGGGAGCCCTATGGAAAAACG 2456

QY 5357 ccagcaacgcggccttttaccggttccctggccttttgcgtggccttttgcctcacatgttct 5416
|||||
Db 2455 CCAGCAACGCGGCCCTTTTACGGTTCCCTGGCCCTTTTGTGGCCCTTTTGTCTCACATGTTCT 2396

QY 5417 ttcctgcgttatccccctgattctctgtggataaacctgattaccgccttttgagtgagctgata 5476
|||||
Db 2395 TTCTCTGCGTTATCCCCCTGATTCTGTGGATAACCCGTATTACC GGCTTTTGTAGTGAGCTGATA 2336

QY 5477 ccgctcgcgcagccgaacgacccgagcgcagcagtcagtgagcgaggaagcggagagc 5536
|||||
Db 2335 CCGCTCGCCGACCGCGAACCAGCCGAGCGCAGGAGTCAGTGAGCGAGGAGCGGAAGAGC 2276

QY 5537 gcccataacgcaaaacccgcctctccccgcggttggccgatttcattaatgcagctggcacg 5596

|||||
Db 2275 GCCCAATACGAAACCGCCTCTCCCGCGCGTGGCCGATTCAATATGCAGCTGGCAGC 2216
QY 5597 acagggttcccgactggaagcgggcagtgagcgcgaacgaattaatgtgagttagctca 5656
|||||
Db 2215 ACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA 2156
QY 5657 ctattaggcacccaggcttacacctttatgtcttcggctcgatgttgtgtggaattg 5716
|||||
Db 2155 CTCAATAGGCACCCAGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAATTG 2096
QY 5717 tgagcgggataacaatttcacacagggaacagctatgacctgattacg 5764
|||||
Db 2095 TGAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACG 2048

RESULT 13
AAA95416
ID AAA95416 standard; DNA; 5733 BP.
XX
AC AAA95416;
XX
DT 12-FEB-2001 (first entry)
XX
DE Aspergillus nidulans pAN52-1 vector.
XX
KW Fusion protein; aqueous two-phase system; ATPS; protein production;
KW protein purification; expression library screening; vector; ds.
XX
OS Aspergillus nidulans.
XX
FH Key Location/Qualifiers
FT promoter 1..2129
FT /*tag= a
FT 2130..2304
FT /*tag= b
FT /product= "gpdA"
FT terminator 2305..3071
FT /*tag= c
XX
PN WO200058342-A1.
XX
PD 05-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-FI00249.
PF
XX 25-MAR-1999; 99FI-0000667.
PR 20-AUG-1999; 99FI-0001782.
XX
PA (VALW) VALTION TEKNILLINEN TUTKIMUSKESKUS.
XX
XX Penttilae M, Nakari-Setaelae T, Fagerstroem R, Selber K, Kula M;
PI Linder M, Tjerneld F;
XX
DR WPI; 2000-686858/67.
XX
PT Isolation and purification of proteins or cells in aqueous two-phase
PT systems, comprises combining a desired protein or a cell to a targeting
PT protein capable of isolating the protein or cell into one of the phases
PT -
XX
PS Example 1; Page 93-95; 109pp; English.
XX
CC The present invention is related to a novel method for separating,
CC purifying and isolating proteins and cells. This involves the use of
CC liquid-liquid extraction in an aqueous two-phase system (ATPS) which
CC partitions molecules by fusing them to targeting proteins which then
CC carry the molecules of interest into one of the phases. The present
CC sequence is an expression vector which was used to demonstrate the method
CC of the invention. The method is useful in also useful in the
CC identification of nucleic acid sequences in expression library screening.
XX
SQ Sequence 5733 BP; 1435 A; 1454 C; 1378 G; 1463 T; 3 other;

Query Match 45.6%; Score 2628; DB 21; Length 5733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3137 cactggccgctgcttttacaaactgctgactgggaaaaaccctggcttacccaaacttaatc 3196
|||||
Db 3095 cactggccgctgcttttacaaactgctgactgggaaaaaccctggcttacccaaacttaatc 3154
QY 3197 gccttgacgacacatcccccttccagctggcgtaatatagcgaagagggccgcacgcgac 3256
|||||
Db 3155 gccttgacgacacatcccccttccagctggcgtaatatagcgaagagggccgcacgcgac 3214
QY 3257 gcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatcgcggtatttctcc 3316
|||||
Db 3215 gcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatcgcggtatttctcc 3274
QY 3317 ttacgcatctgtcggtatttcacacccgcacatattggtgcactctcagtacaaatctgctctg 3376
|||||
Db 3275 ttacgcatctgtcggtatttcacacccgcacatattggtgcactctcagtacaaatctgctctg 3334
QY 3377 atgcgcgcatagttaagccagcccgacaccccgcaacaccccgctgcagcgcctgacggg 3436
|||||
Db 3335 atgcgcgcatagttaagccagcccgacaccccgcaacaccccgctgcagcgcctgacggg 3394
QY 3437 cttgtctgctcccgcatccgcttacagacaaagctgtgacctctccgggagctgcattg 3496
|||||
Db 3395 cttgtctgctcccgcatccgcttacagacaaagctgtgacctctccgggagctgcattg 3454
QY 3497 gtcagagggttttcacccgtcatcacggaaaacgcgcgagacgaaaggccctcgtgatacgc 3556
|||||
Db 3455 gtcagagggttttcacccgtcatcacggaaaacgcgcgagacgaaaggccctcgtgatacgc 3514
QY 3557 tatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcaccttttc 3616
|||||
Db 3515 tatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcaccttttc 3574
QY 3617 ggggaaatgtgcgcggaaacccctatttggtttatttttctaataacattcaaatatgtatc 3676
|||||
Db 3575 ggggaaatgtgcgcggaaacccctatttggtttatttttctaataacattcaaatatgtatc 3634
QY 3677 cgctcatgagacaataaacctgataaaatgcttcaataatatattgaaaaaggaagagtatga 3736
|||||
Db 3635 cgctcatgagacaataaacctgataaaatgcttcaataatatattgaaaaaggaagagtatga 3694
QY 3737 gtattcaaatctccgtgtcgcccttattcccttttttgcggcattttgccttccctgttt 3796
|||||
Db 3695 gtattcaaatctccgtgtcgcccttattcccttttttgcggcattttgccttccctgttt 3754
QY 3797 ttgctcacccagaaaacgctggtgaaagtataaagatgctgaagatcagttgggtgcacgag 3856
|||||
Db 3755 ttgctcacccagaaaacgctggtgaaagtataaagatgctgaagatcagttgggtgcacgag 3814
QY 3857 tgggttacatcgaaactggatctcaacagcgggtaagatcccttgagagttttcgccccgaag 3916
|||||
Db 3815 tgggttacatcgaaactggatctcaacagcgggtaagatcccttgagagttttcgccccgaag 3874
QY 3917 aacgttttccaatgatgagcacacttttaagttctgctatgtggcgcggtattatcccgta 3976
|||||
Db 3875 aacgttttccaatgatgagcacacttttaagttctgctatgtggcgcggtattatcccgta 3934
QY 3977 ttgacgcggggcaagagcaactcggtcgcgcacatacacactattctcagaatgacttggtg 4036
|||||
Db 3935 ttgacgcggggcaagagcaactcggtcgcgcacatacacactattctcagaatgacttggtg 3994
QY 4037 agtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgca 4096
|||||
Db 3995 agtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgca 4054
QY 4097 gtgctgccaataaccatgagtataaacactgcggccaacttacttcttgacaacgacgag 4156
|||||
Db 4055 gtgctgccaataaccatgagtataaacactgcggccaacttacttcttgacaacgacgag 4114

QY 4157 gaccgaaggagcctaaccgcttttttgcacacaatgggggatcatgttaactgccttgatc 4216
|||||
Db 4115 gaccgaaggagcctaaccgcttttttgcacacaatgggggatcatgttaactgccttgatc 4174

QY 4217 gttgggaaccggagctgaatgaagccataccaaaacgacgagcgtgacaccacgatgcctg 4276
|||||
Db 4175 gttgggaaccggagctgaatgaagccataccaaaacgacgagcgtgacaccacgatgcctg 4234

QY 4277 tagcaatggcaacaacgctgcgcaaaactattaaactggcgaactacttactctagcttccc 4336
|||||
Db 4235 tagcaatggcaacaacgctgcgcaaaactattaaactggcgaactacttactctagcttccc 4294

QY 4337 ggcaacaattaatagactggatggaggcgagataaaagtgcaggaccacttctgcgctcgg 4396
|||||
Db 4295 ggcaacaattaatagactggatggaggcgagataaaagtgcaggaccacttctgcgctcgg 4354

QY 4397 cccttcgggctggctggtttattgtctgataaatctggagccctcccgtatcgttagttatctacacga 4516
|||||
Db 4355 cccttcgggctggctggtttattgtctgataaatctggagccctcccgtatcgttagttatctacacga 4474

QY 4457 gtatcattgcagcactggggccagatggttaagccctcccgtatcgttagttatctacacga 4516
|||||
Db 4415 gtatcattgcagcactggggccagatggttaagccctcccgtatcgttagttatctacacga 4474

QY 4517 cggggagtcagggaactatggatgaacgaataagacagatcgctgagataggtgcctcac 4576
|||||
Db 4475 cggggagtcagggaactatggatgaacgaataagacagatcgctgagataggtgcctcac 4534

QY 4577 tgattaaagcattggtaactgtcagaccaaagtttactcatatatacttttagattgatttaa 4636
|||||
Db 4535 tgattaaagcattggtaactgtcagaccaaagtttactcatatatacttttagattgatttaa 4594

QY 4637 aacttcaatttttaatttaaaaggatctaggtgaagatcccttttttgataaatctcatgacca 4696
|||||
Db 4595 aacttcaatttttaatttaaaaggatctaggtgaagatcccttttttgataaatctcatgacca 4654

QY 4697 aaatcccttaacgtgagttttctgttccactgagcgtcagaccccccttagaaaaagatcaaaag 4756
|||||
Db 4655 aaatcccttaacgtgagttttctgttccactgagcgtcagaccccccttagaaaaagatcaaaag 4714

QY 4757 gatcttcttgagatcccttttttctgcgcgtaaatctgctgcttgcaacaaaaaaaccac 4816
|||||
Db 4715 gatcttcttgagatcccttttttctgcgcgtaaatctgctgcttgcaacaaaaaaaccac 4774

QY 4817 cgctaccagcgtggtgtgttgcgggatacaagagctaccaaactcttttccgaaggtaa 4876
|||||
Db 4775 cgctaccagcgtggtgtgttgcgggatacaagagctaccaaactcttttccgaaggtaa 4834

QY 4877 ctggcttcagcagagcgcagataccaaaatactgtcccttctagtgtagccgtagttaggcc 4936
|||||
Db 4835 ctggcttcagcagagcgcagataccaaaatactgtcccttctagtgtagccgtagttaggcc 4894

QY 4937 accacttcaagaactctgttagcaccgcctacataacctcgctctgctgctaactcctgttaccag 4996
|||||
Db 4895 accacttcaagaactctgttagcaccgcctacataacctcgctctgctgctaactcctgttaccag 4954

QY 4997 tggctgctgcagtgccgataaagtctgttaccgggttggaactcaagacgatagttac 5056
|||||
Db 4955 tggctgctgcagtgccgataaagtctgttaccgggttggaactcaagacgatagttac 5014

QY 5057 cggataaggcgagcggtcgggctgaacggggggttcgtgcacacgcccagcttggagc 5116
|||||
Db 5015 cggataaggcgagcggtcgggctgaacggggggttcgtgcacacgcccagcttggagc 5074

QY 5117 gaacgacctacaccgaactgagataacctacacgctgagctatgagaaagcgccacgcttc 5176
|||||
Db 5075 gaacgacctacaccgaactgagataacctacacgctgagctatgagaaagcgccacgcttc 5134

QY 5177 ccgaaggggagaaaaggcggacaggtatccgggtaagcggcaggggtcggaacaggagagcgca 5236
|||||
Db 5135 ccgaaggggagaaaaggcggacaggtatccgggtaagcggcaggggtcggaaacaggagagcgca 5194

QY 5237 cgagggagcttccaggggggaaaacgcctggtatctttatagtcctgctcgggttccgccacc 5296
|||||
Db 5195 cgagggagcttccaggggggaaaacgcctggtatctttatagtcctgctcgggttccgccacc 5254

QY 5297 tctgacttgagcgtcgatttttgtgtgatcgtcgtcagggggcgagcctatggaaaaaacg 5356
|||||
Db 5255 tctgacttgagcgtcgatttttgtgtgatcgtcgtcagggggcgagcctatggaaaaaacg 5314

QY 5357 ccagcaacgcgccttttttacggttccctggttccctggttctggtggccttttgcacatgttct 5416
|||||
Db 5315 ccagcaacgcgccttttttacggttccctggttccctggttctggtggccttttgcacatgttct 5374

QY 5417 ttctgctgttatccctgatctctgtggataaccgtattaccgcctttgagtgagctgata 5476
|||||
Db 5375 ttctgctgttatccctgatctctgtggataaccgtattaccgcctttgagtgagctgata 5434

QY 5477 ccgctcgcgcagccgaacgacgagcgcgagtcagtgagcgaagcgaagagc 5536
|||||
Db 5435 ccgctcgcgcagccgaacgacgagcgcgagtcagtgagcgaagcgaagagc 5494

QY 5537 gcccaatacgcaaaacccgctctctcccgcgcgttgccgattcattaatgcagctggcacg 5596
|||||
Db 5495 gcccaatacgcaaaacccgctctctcccgcgcgttgccgattcattaatgcagctggcacg 5554

QY 5597 acaggtttccgcactggaaaagcgggcagtcgagcgaacgaataatgtgagttagctca 5656
|||||
Db 5555 acaggtttccgcactggaaaagcgggcagtcgagcgaacgaataatgtgagttagctca 5614

QY 5657 ctcataggcaccgccaggtttacactttatgcttccggtcgtatgttgtggaattg 5716
|||||
Db 5615 ctcataggcaccgccaggtttacactttatgcttccggtcgtatgttgtggaattg 5674

QY 5717 tgagcgggataacaatttcacacagggaaaacagctatgaccatgattacg 5764
|||||
Db 5675 tgagcgggataacaatttcacacagggaaaacagctatgaccatgattacg 5722

RESULT 14
AAx82259/c
ID AAX82259 standard; cDNA; 9164 BP.
XX
AC AAX82259;
XX
DT 18-AUG-1999 (first entry)
XX
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).
XX
KW Factor VIII protein; gene modification; gene therapy; clinical disorder;
KW splicing pattern; RNA processing; gene regulation; beta-domain; human;
KW ss.
XX
OS Homo sapiens.
XX
PN WO929848-A1.
XX
PD 17-JUN-1999.
XX
PF 25-NOV-1998; 98WO-US25354.
XX
PR 16-JAN-1998; 98US-0071596.
PR 05-DEC-1997; 97US-0067614.
XX
PA (IMMU-) IMMUNE RESPONSE CORP.
XX
PI Bidlingmaier S, Gonzales JEN, Ill CR, Yang CQ;
XX
DR WPI; 1999-385602/32.
DR P-PSDB; AAY21675.
XX
PT Genes and vectors exhibiting increased expression and novel splicing
PT patterns, useful for expression of, e.g. beta-domain deleted factor
PT VIII
XX

QY 4937 accacttcaagaactctgtagcacgcgcctacatacctcgctctgctaatacctgttaccag 4996
|||||
Db 7347 ACCACTTCAGAAGTCTGTAGCACCGCCTACATACCTCGCTCTGCTAACTCCTGTACCAG 7288

QY 4997 tggctgctgccagtggcgataagtcgtgtcttaccgggttgacctcaagacgatatgttac 5056
|||||
Db 7287 TGGCTGCTGCCAGTGGCGATAAGTCGTCTTTACCGGGTTGGACTCAAGACGATAGTTAC 7228

QY 5057 cggataaaggcgagcggctcgggctgaacggggggttcgtgcacacagcccgcttgagc 5116
|||||
Db 7227 CGGATAAGGCGCAGCGGTGGGGCTGAACGGGGGGGTTCTGTGCACACAGCCACGCTTGAGC 7168

QY 5117 gaacgacctacacccgaactgagatatacctacagcgtgagctatgagaagcgccacgcttc 5176
|||||
Db 7167 GAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAGCGCCACGCTTC 7108

QY 5177 ccgaaggggagaaaggcggacaggtatccgggtaagcggcagggctcggaacagagagcgca 5236
|||||
Db 7107 CCGAAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGCGCA 7048

QY 5237 cgagggagctccaggggggaaacgcctgggtatctttatagtcctctcgggtttcgccacc 5296
|||||
Db 7047 CGAGGGAGCTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACC 6988

QY 5297 tctgacttgagcgtcgattttttgtgctgctcgtcagggggggggagcctatggaaaaaacg 5356
|||||
Db 6987 TCTGACTTGAGCGTCGATTTTGTGTGATGCTCGTCAGGGGGGGGAGCCTATGGAAAAACG 6928

QY 5357 ccagcaacggcgcccttttttacggttcctggcccttttgcggccttttgcacatgttct 5416
|||||
Db 6927 CCAGCAACGGCGCCCTTTTACGGTTCTTGCCCTTTTGTGCGCCTTTTGCTCACATGTTCT 6868

QY 5417 ttctgcgttatccccctgattctgttgataaccgtattaccgcctttgagtgagctgata 5476
|||||
Db 6867 TTCTCTGCTTATCCCTGTATTCTGTGATAACCGTATTACCGCCTTTGAGTGAGCTGATA 6808

QY 5477 ccgctcgcgcagccggaacgacgcagcgagcagtcagtcagtcagtcagtcagtcagtcagtc 5536
|||||
Db 6807 CCGCTCGCGCAGCGCGAACGACGACGAGCGAGCGAGTCAGTGAGCGAGGAAGCGAAGAGC 6748

QY 5537 gcccaataacgcaaacccctctccccgcgcttgccgattcattcaatgcagctggcacg 5596
|||||
Db 6747 GCCCAATACGAAACCGCCTCTCCCCGCGCTTGCCCGATTTCATTAAATGCAGCTGGCACC 6688

QY 5597 acaggtttcccgactggaaagcggcagtcagtcagtcagtcagtcagtcagtcagtcagtc 5656
|||||
Db 6687 ACAGGTTTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA 6628

QY 5657 ctcattaggcaacccagcgttttacactttatgcttccggctcgtatgctgtgtggaattg 5716
|||||
Db 6627 CTCATTAGGCACCCACGGCTTTTACACTTATGCTTCCGGCTCGTATGTTGTGTGAATTG 6568

QY 5717 tgagcggataacaatttcacacaggaacagctatgaccatgattacg 5764
|||||
Db 6567 TGAGCGGATAACAATTTACACACAGGAAACAGCTATGACCATGATTACG 6520

RESULT 15
AAX04251/c
ID AAX04251 standard; DNA; 9632 BP.
XX
AC AAX04251;
XX
DT 15-APR-1999 (first entry)
XX
DE Baculovirus transfer vector pVLI393 DNA sequence.
XX
KW Ricin-like toxin; cancer; viral infection; parasitic infection;
KW linker; B chain; A chain; protease; fungal infection; malaria;
KW leucocyte proliferation; cytomegalovirus; herpes; hepatitis;
KW rhinovirus; laryngeotracheitis; poliomyelitis; varicella zoster;
KW cystic fibrosis; multiple sclerosis; ds.

.

OS Escherichia coli.
OS Synthetic.
XX
PN WO9849311-A2.
XX
PD 05-NOV-1998.
XX
PF 30-APR-1998; 98WO-CA00394.
XX
PR 29-OCT-1997; 97US-0063715.
PR 30-APR-1997; 97US-0045148.
XX
PA (DNOV-) DE NOVO ENZYME CORP.
XX
PI Borgford T;
XX
DR WPI; 1999-009431/01.
XX
PT New nucleic acid encoding ricin-like toxin with an interchain linker
PT cleaved by protease - is specific for diseased cells, useful for,
PT e.g. killing selectively cancer or infected cells
XX
PS Disclosure; Fig 1; 352pp; English.
XX
CC The present invention describes new purified and isolated nucleic acids
CC (I) encoding: (i) the A and B chains of a ricin-like toxin (II); and
CC (ii) a heterologous linker, joining the two chains and including a
CC cleavage recognition site for a disease-specific protease (III). Also
CC described are: (1) plasmids or baculovirus transfer vectors that contain
CC (I); and (2) recombinant protein (IV) consisting of the A and B chains
CC of (II) joined by the specified linker. (IV), produced by expression of
CC (I) in host cells, are used to inhibit or kill diseased cells that
CC produce (III), particularly for treating cancers (e.g. leucocyte
CC proliferation; cancer of ovary, pancreas, breast or prostate; glioma) or
CC infections caused by fungi, parasites (e.g. malaria) or viruses (e.g.
CC cytomegalovirus (CMV), herpes, hepatitis, rhinovirus, laryngeotracheitis,
CC poliomyelitis or varicella zoster), also cystic fibrosis and multiple
CC sclerosis. Alternatively, (I) is used to express (IV) in vivo. (IV) is
CC toxic specifically for (III)-expressing cells and does not depend for
CC specificity on a cell-binding component. When used to treat virus-
CC infected cells, transcytosis and cytotoxicity of (IV) are increased by
CC retrograde translocation from endoplasmic reticulum to cytoplasm (which
CC some viruses exploit to avoid immune detection), so selectivity and
CC safety are further improved. (IV) are not toxic until chain A is
CC released and this occurs only in target cells. The present sequence
CC represents a nucleotide sequence from the present invention.
XX
SQ Sequence 9632 BP; 2602 A; 2122 C; 2176 G; 2732 T; 0 other;

Query Match 45.6%; Score 2628; DB 20; Length 9632;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3137 cactggcgcgtgttttacaacgctgctgactgggaaacccctggcgttaccacaacttaatc 3196
|||||
Db 9630 CACTGGCCGTCGTTTACAAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAACCTAATC 9571

QY 3197 gccttgcaacacatcccccttttcgccagctggcgtaataagcgaaagggcccgaccgac 3256
|||||
Db 9570 GCCTTGACACACATCCCCCTTTTCGCCACGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC 9511

QY 3257 gcccttcccaacagattgcgcagcctgaatggcgaaatggcgccctgatgcgggtatttctcc 3316
|||||
Db 9510 GCCCTTCCCAACAGTTGGCGAGCCCTGAATGGCGAATGGCGCCTGATGCGGTATTTCTCC 9451

QY 3317 ttacgcattctgcgggtatttcacacgcgcataatggtgcactctcagtaacaatctgctctg 3376
|||||
Db 9450 TTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACCTCTCAGTACAAATCTGCTCTG 9391

QY 3377 atgccgcataagtaagccagccccgcacccccgcacccccgcacccccgcacccccgcacggg 3436
|||||
Db 9390 ATGCCCGCATAGTTAAGCCAGCCCCCGACACCCCGCAACCCCGCTGACGCGCCTTGACGGG 9331

QY 3437 cttgtctgctcccgccatccgcttacagacaagctgtgaccgtctccgggagctgcatgt 3496
|||||
Db 9330 CTTGTCTGCTCCGGCATCCGGTTACAGACAAGCTGTGACCGCTCTCCGGGAGCTGCATGT 9271

QY 3497 gtcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaaaaggcctcgtgatacgcc 3556
|||||
Db 9270 GTCAGAGGTTTTCACCGTCAATACCGAAACGCGGAGACGAAAGGGCCTCGTGATACGCC 9211

QY 3557 tattttataggttaatgtcatgataataatggtttcttagacgtcaagtgccacttttc 3616
|||||
Db 9210 TATTTTATAGGTTAATGTCATGATAAATAGGTTTCTTAGACGTCAAGTGGCACTTTTC 9151

QY 3617 ggggaaatgtgcgcggaacccctatttgttatttttctaaatacatcattcaaatatgtatc 3676
|||||
Db 9150 GGGGAAATGTGCGCGGAACCCCTATTGTGTTATTTTCTAAATACATTCAAAATATGTATC 9091

QY 3677 cgctcatgagacaataaacctgataaaatgccttcaataataattgaaaggaagagatatga 3736
|||||
Db 9090 CGCTCATGAGACAATAACCCTGATAAAATGCTTCAATAATATTGAAAAAGGAAGATATGA 9031

QY 3737 gtattcaacatttcctgctgctgccttatttcccttttttgcggcattttgccttcctgttt 3796
|||||
Db 9030 GTATTCAACATTTCGGTGTCCGCCTTATTCCCTTTTGTGGCCATTGTGCCTTCTGTGTTT 8971

QY 3797 ttgctcaccgagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgag 3856
|||||
Db 8970 TTGCTCACCGAAGAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 8911

QY 3857 tgggttacatgaactggatctcacaacggtgaagatcccttgagagtttttcgccccgaag 3916
|||||
Db 8910 TGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAG 8851

QY 3917 aacgtttttccaatgatgagcaccttttaagttctgctatgtggcgcggtattatcccgta 3976
|||||
Db 8850 AACGTTTTCGAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTA 8791

QY 3977 ttgacgcgcggcaagagcaactcggctgcgcgcatacactattctcagaatgacttggtg 4036
|||||
Db 8790 TTGACGCCGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTG 8731

QY 4037 agtactcacagtcacagaaaaagcatcttacggatggcatggcatgacagtaagagaattatgca 4096
|||||
Db 8730 AGTACTCACAGTCACAGAAAAGCATCTACGGATGGCATGACAGTAAGAGAATATATGCA 8671

QY 4097 gtgctgccataaaccatgagtataaacactgcggccaaacttacttctgacaacgatcggag 4156
|||||
Db 8670 GTGCTGCCATAAACCATGAGTGATAACACTCGGCCAACTTACTTCTGACAACGATCGGAG 8611

QY 4157 gaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgcttgatc 4216
|||||
Db 8610 GACCGAAGGAGCTAACCGCTTTTTCACAACATGGGGATCATGTAACTCGCCTTGATC 8551

QY 4217 gttgggaacccggagctgaatgaagccataaccaaacgacgagcgtgacaccacgatgcctg 4276
|||||
Db 8550 GTTGGGAACCCGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACCAGATGCCTG 8491

QY 4277 tagcaatggcaacaacgttgcgcaaaactattaactggcgaaactacttactctagcttccc 4336
|||||
Db 8490 TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC 8431

QY 4337 ggcaacaattaatagactggatggagggcgataaaagtgcaggaccacttctgcgctcg 4396
|||||
Db 8430 GGCNACAATTAAATAGACTGGATGGAGCGCGGATAAAGTTGCAGGACCACCTCTGCGCTCGG 8371

QY 4397 cccttcggctggctggtttatttgctgataaaatctggagccggtgagcgtgggtctcgcg 4456
|||||
Db 8370 CCTTCCGGCTGGCTGGTTTATTGCTGATAAACTCTGGAGCCGGTGAGCGGTGGGTCTCGCG 8311

QY 4457 gtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacga 4516
|||||
Db 8310 GTATCATTCAGCACATGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGA 8251

QY 4517 cggggagtcaggcaactatggatgaacgaatagacagatcgctgagataggtgcctcac 4576
|||||
Db 8250 CGGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTGCCTCAC 8191

QY 4577 tgattaagcatgtgtaactgtcagaccaaagtttactcatatatatacttttagattgattaa 4636
|||||
Db 8190 TGATTAAGCATTTGGTAACGTGTCAAGCCAAGTTACTCATATATACTTTTAGATTTGATTAA 8131

QY 4637 aacttcaatttttaatttaaaaggatctagtggaagatcccttttggataatctcatgacca 4696
|||||
Db 8130 AACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATGACCA 8071

QY 4697 aaatcccttaacgtgagttttcgttccactgagcgtcagacccccctagaaaaagatcaaa 4756
|||||
Db 8070 AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAGATCAAAAG 8011

QY 4757 gatcttcttgagatcccttttttctgcgcgtaactctgctgctgcaaaaaaaaccac 4816
|||||
Db 8010 GATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGTGCAAAACAAAAAACCCAC 7951

QY 4817 cgctaccagcgggtggtttgttgcgggatcaagagctaccactcttttccgaaggtaa 4876
|||||
Db 7950 CGCTACCAAGCGGTGTTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA 7891

QY 4877 ctggcttcagcagcgcagataccaaaatactgtccttcttagtgtagccgtagttaggcc 4936
|||||
Db 7890 CTGGCTTCAGCAGAGCGCAGATACCAAATAC1GTCTCTTAGTGTAGCCGTAGTAGGCC 7831

QY 4937 accacttcaagaaactctgttagcaccgcctacatacctcgtcgtctgtaactcctgttaccag 4996
|||||
Db 7830 ACCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCCTGTACCAG 7771

QY 4997 tggctgctgccagtggcgataaagtcgtgtcttaccgggttgactcaagacgatatgttac 5056
|||||
Db 7770 TGGCTGCTGCCAGTGGCGATAAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTAC 7711

QY 5057 cggataaggcgcgcggtcggtgaaacgggggtcggtgcacacagccccagcttggagc 5116
|||||
Db 7710 CGGATAAGCGCGCAGCGGTGGGTGAACGGGGGTTCGTGCACACAGCCCGCTTGGAGC 7651

QY 5117 gaacgacctacacccaactgagatatacctacagcgtgagctatgagaaaagcgcacgcttc 5176
|||||
Db 7650 GAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCCGCTTTC 7591

QY 5177 ccgaaggagaaaaggcggacaggtatccggtaacggcagggcagggcgaacaggaagcgcga 5236
|||||
Db 7590 CCGAAGGGAGAAAAGCGCGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGCGCA 7531

QY 5237 cgaggagcttccagggggaaacgcctggtatctttatagctcgtcggttccgcccac 5296
|||||
Db 7530 CGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTATAGTCTCTGTCGGGTTTCGCCACC 7471

QY 5297 tctgacttgagcgtcgatttttgatgctcgtcagggggggcgagcctatggaaaaacg 5356
|||||
Db 7470 TCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGAAAAACG 7411

QY 5357 ccagcaacgcggcctttttacggttccctggccttttgccttttgcacacatgttct 5416
|||||
Db 7410 CCAGCAACGCGGCCCTTTTACGGTTCCCTGGCCCTTTTGCTGGCCCTTTTGCTCACATGTTCT 7351

QY 5417 ttccctgcgttatccctgattctgtggataaacggtattaccgccttttgagtgagctgata 5476
|||||
Db 7350 TTCTCTGCGTTATCCCTCGATTCTGTGGATAACCCGTATTACCGCCTTTTGTAGTGAGCTGATA 7291

QY 5477 ccgctcgcgcagccgaacgacgcagcgcagcagtcagtcagtcagtcagtcagtcagtcagtc 5536
|||||
Db 7290 CCGCTCGCGCAGCGCAACGACCGAGCGCAGCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG 7231

QY 5537 gcccaatacgcaaacgcctctccccgcgcgttgccgattcattaaatgcagctggcgacg 5596
|||||
Db 7230 GCCCAATACGCAAAACCGCCTCTCCCCCGCGCTTGGCCGATTCTTAATGCAGCTGGCAGC 7171

QY 5597 acaggtttcccgactggaaaaagcggcagtcgagcgcaaacgcaaatatgtgagttagctca 5656

```
Db 7170 ACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAAATGTAGTCTCA 7111
Qy 5657 ctcataggcaccgccgctttacactttatgcttccggctcgatgtgtggaattg 5716
Db 7110 CTCATTAGSCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTGTGGAATTG 7051
Qy 5717 tgagcggataacaatttcacacagggaacagctatgaccatgattacg 5764
Db 7050 TGAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGATTACG 7003
```

Search completed: August 31, 2002, 22:58:22
Job time: 16803 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 19:58:51 ; Search time 6064.54 Seconds
(without alignments)
12834.764 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcatgctgcaggtc.....ctatgaccatgattacgcca 5767

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_gss:*
 - 13: em_gss_hum:*
 - 14: em_gss_inv:*
 - 15: em_gss_pln:*
 - 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	18.4	2080	11 BC001541	BC001541 Homo sapi
2	864	15.0	1070	9 AJ281552	AJ281552 4A3A-P6F1
3	841	14.6	841	9 AL042026	AL042026 DKFZp434E
4	835	14.5	1067	9 AU081137	AU081137 AU081137
5	828	14.4	1089	9 AU081124	AU081124 AU081124
6	723	12.5	954	9 AL044364	AL044364 DKFZp434C
7	723	12.5	1004	9 AJ281480	AJ281480 4A3A-P4G8
8	691	12.0	1013	10 BM438846	BM438846 IPLVr0015
9	691	12.0	1039	9 AU081040	AU081040 AU081040
10	686	11.9	780	10 BI753192	BI753192 603026066
11	682	11.8	1163	9 AU081044	AU081044 AU081044
12	679	11.8	680	9 AL646751	AL646751 AL646751
13	678	11.8	707	9 AL656688	AL656688 AL656688
14	678	11.8	800	9 AJ281449	AJ281449 4A3A-P4D5
15	673	11.7	724	9 AL645114	AL645114 AL645114
16	672	11.7	715	9 AL661706	AL661706 AL661706
17	662	11.5	705	9 AL643164	AL643164 AL643164

C 18	659	11.4	659	9	AL643220	AL643220	AL643220	AL643220
C 19	659	11.4	689	9	AL646532	AL646532	AL646532	AL646532
C 20	657	11.4	847	12	AZ687169	AZ687169	ENTMP05TF	AZ687169
C 21	652	11.3	653	9	AL662063	AL662063	AL662063	AL662063
C 22	648	11.2	648	9	AL640650	AL640650	AL640650	AL640650
C 23	648	11.2	703	9	AJ281437	AJ281437	4A3A-P4C3	AJ281437
C 24	647	11.2	767	9	AL040542	AL040542	DKFZp434I	AL040542
C 25	645	11.2	645	9	AL639182	AL639182	AL639182	AL639182
C 26	645	11.2	670	9	AL662130	AL662130	AL662130	AL662130
C 27	645	11.2	717	12	AZ208376	AZ208376	SP_0150_A	AZ208376
C 28	642	11.1	642	9	AL642207	AL642207	AL642207	AL642207
C 29	642	11.1	670	9	AL635952	AL635952	AL635952	AL635952
C 30	640	11.1	640	12	B84895	B84895	RPC111-28L6	B84895
C 31	638	11.1	832	10	BG923768	BG923768	602825893	BG923768
C 32	636	11.0	687	9	AU001481	AU001481	AU001481	AU001481
C 33	633	11.0	688	9	AL042640	AL042640	DKFZp434M	AL042640
C 34	629	10.9	657	9	AL641508	AL641508	AL641508	AL641508
C 35	628	10.9	669	9	AL660789	AL660789	AL660789	AL660789
C 36	627	10.9	705	9	AL635845	AL635845	AL635845	AL635845
C 37	624	10.8	637	9	AL627526	AL627526	AL627526	AL627526
C 38	623	10.8	675	9	AL636713	AL636713	AL636713	AL636713
C 39	621	10.8	647	9	AL642844	AL642844	AL642844	AL642844
C 40	619	10.7	645	9	AL642835	AL642835	AL642835	AL642835
C 41	619	10.7	655	9	AL659614	AL659614	AL659614	AL659614
C 42	619	10.7	664	9	AL639797	AL639797	AL639797	AL639797
C 43	617	10.7	650	9	AL638248	AL638248	AL638248	AL638248
C 44	613	10.6	672	9	AL640799	AL640799	AL640799	AL640799
C 45	612	10.6	663	9	AU001472	AU001472	AU001472	AU001472

ALIGNMENTS

RESULT 1

BC001541
LOCUS BC001541 Homo sapiens, clone IMAGE:3453362, mRNA. 2080 bp linear HTC 31-JAN-2002
DEFINITION BC001541
ACCESSION BC001541
VERSION BC001541.1 GI:14705895
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2080)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 4 Row: e Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7710156
This clone has the following problem: no polyA-tail.

FEATURES

source

1. .2080
/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3453362"

/tissue_type="Placenta, choriocarcinoma"

/clone_lib="NIH_MGC_10"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

BASE COUNT

319 a 692 c 651 g 418 t

ORIGIN

Query Match

18.4%; Score 1064; DB 11; Length 2080;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

844

ccccgcagtgctgctgcacacgccttcctcgcttcctcccaactcttctcctcctc

903

Db

104

CCCCCGCAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCACACTCTTCTCCTCCTCCTC

163

QY

904

tggctcctggggtggaggagtggggctgagggccgggagggatgcagagctgctggtgacg

963

Db

164

TGGCTCCTGGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTGGTGACG

223

QY

964

gtcgtggggccggctgcggyggcattgcgcctgaagacccccggggccctgtctctgct

1023

Db

224

GTGCGTGGGGCCGGCTGCGGGGCATTCGCCTGAAGACCCCCGGGGCCCTGTCTCTGTCT

283

QY

1024

ttcctgggcaccccttgcggagccaccccatgggacccccctgcgttcttctgccaccggag

1083

Db

284

TTCCCTGGGCATCCCCCTTTCGGGAGCCACCCCATGGGACCCCGTCGCTTCTGCCACCGGAG

343

QY

1084

cccaagcagccttggtcaggggtgtagacgctacaaccttccagagtgtctgtctaccaa

1143

Db

344

CCCAAGCAGCCTTGGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGTACTACAA

403

QY

1144

tatgtggacacccctataccagggttttgagggcaccgagatgtggaaccccccaaccgtgag

1203

Db

404

TATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAACCGTGAG

463

QY

1204

ctgagcaggagactgcctgtacctcaactgtgtggacaccataccccggcctacatcccc

1263

Db

464

CTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGCGCCTACATCCCCC

523

QY

1264

acccctgcctcgtctggtatctatgggggttgcttctacagtggggcctcctcttgga

1323

Db

524

ACCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCCTTGGAC

583

QY

1324

gtgtacgatggccgcttcttggttacaggccgagagactgtgctggttccatgaactac

1383

Db

584

GTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGTGTCCATGAACATC

643

QY

1384

cgggtgggagcctttggcttcctggccctgcgggagccgagagggccccgggcaatgtg

1443

Db

644

CGGGTGGAGACCTTTGGCTTCTTGGCCCTGCCGGGAGCCGAGAGGCCCGGGCAATGTG

703

QY

1444

ggtcctcctggatcagaggctggccctgcagtgggtgcaggagaaactggtggccttcggg

1503

Db

704

GGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCCTTCGGG

763

QY

1504

ggtgacccgacatcagtgacgctgttttggggagagcgcgggagccctcggtgggcatg

1563

Db

764

GGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTGGGCATG

823

QY

1564

cacctgctgtccccgccagccgggacctgttccacagggccgtgctgcagagcggtgcc

1623

Db

824

CACCTGCTGTCCCCGCCAGCCGGGCCCTGTTCCACAGGGCCGTGTCTGCAGAGCGGTGCC

883

QY

1624

cccaatggacccctgggccacggtgggcatgggagagggccctgcagggccacgcagctg

1683

Db

884

CCCAATGGAACCTTGGGCCACGCTGGGCATGGGAGAGGCCCGCTCGCAGGGCCACGCAGCTG

943

QY

1684

gcccacctgtgggctgtcctccaggcggcactggtgggaatgacacagagctggtagcc

1743

Db

944

GCCCACTTGTGGGCTGTCTCCAGCGCGGCACGTGGTGGGAATGACACAGAGCTGGTAGCC

1003

Qy	1744	tgccttcggacacgaccagcgaggctcctggtgaaccacgaatggcacgtgctgcctcaa	1803	
Db	1004	TGCCTTCGGACACGACGACGCGCAGGTCTCGTGAACCAACCAAGATGGCACGTGCTGCCTCAA	1063	
Qy	1804	gaaagcgtcttcgggttctccttcgtgcctgtggtagatggagacttctcagtgacacc	1863	
Db	1064	GAAAGCGTCTTCGGGTCTCTCCTTCGTGCCTGTGGTAGATGGAGACTTCCTCAGTGACACC	1123	
Qy	1864	ccagaggccctcatcaacgcgggagacttccacggcctgcaggt	1907	
Db	1124	CCAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGT	1167	
RESULT 2				
AJ281552/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
COMMENT				
FEATURES				
source				
BASE COUNT				
ORIGIN				

Query Match 15.0%; Score 864; DB 9; Length 1070;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 964; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
Qy	4398	ccttcggctggtggtttattgtgataaaatctggagccggtgagcgtggtctcgcg	4457						
Db	966	CCTTCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGGGTCTCGCGG	907						
Qy	4458	tatcattgcagcactggggccagatggttaagccctcccgatatgtagttatctacacgac	4517						
Db	906	TATCATTCAGCACACTGGGGCCAGATGGTAAGCCCTCCCGTATCTAGTATCTACACGAC	847						

Qy	4518	ggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcact	4577
Db	846	GGGAGTCAGGCAACTATGGATGGACGAAATAGACAGATCGCTGAGATAGTGCCTCACT	787
Qy	4578	gattaagcattggttaactgtcagaccaagtttactcatatatacttacttagattgattaaa	4637
Db	786	GATTAAGCATTGGTAACGTGCAGACCAAGTTTACTCATATATACTTTAGATTGATTAAA	727
Qy	4638	acttcatttttaatttaaaggatctagggtgaagatccctttttgataaatctcatgaccaa	4697
Db	726	ACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCITTTTIGATAATCTCATGACCAA	667
Qy	4698	aatcccttaacgtgagtttttcgttccactgagcgtcagaccccgtagaaaaagatcaaa	4757
Db	666	AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAA	607
Qy	4758	atctcttgagatccctttttctgcgcgtaactcgtcgtcgcacaaacaaaaccacc	4817
Db	606	ATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTGCAACCAAAAAACCACC	547
Qy	4818	gctaccagcgggtgttgttgcgggatcaagagctacacactcttttccgaaggtaac	4877
Db	546	GCTACCAGCGGTGTTGTTTGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC	487
Qy	4878	tggcttcagcagcgcagatcaccaataactgtccttctagtgtagccgtagttaggcca	4937
Db	486	TGSGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGTGTAGCCGTAGTAGGCCA	427
Qy	4938	ccacttcaagaactctgtagcaccgcctacataacctcgtcgtctaactcctgtttaccagt	4997
Db	426	CCACTTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCTGTCTAATCCTGTTACCAGT	367
Qy	4998	ggctgctgccagtgccgataaagtcgtgtcttaccgggttggaactcaagacgatagttacc	5057
Db	366	GGCTGCTGCCAGTGGCGATAAGTCGTGCTTTACCGGGTTGGACTCAAGACGATAGTTACC	307
Qy	5058	ggataaggcgcagcgggtcggtgaacgggggttcgtgcacacagcccagcttggagcg	5117
Db	306	GGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTTCGTGCACACAGCCAGCTTGGAGCG	247
Qy	5118	aacgacctacacccgaactgagatacctacacagcgtgagctatgagaaagcgccacgttcc	5177
Db	246	AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCC	187
Qy	5178	cgaaggagaaaaggcggacaggtatccggtaagcggcagggtcggaaacaggagagcgcac	5237
Db	186	CGAAGGAGAAAGCGGCACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCAC	127
Qy	5238	gagggagcttccagggggaaacgcctggtatctttatagtcctgctcgggtttcgccacct	5297
Db	126	GAGGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTGTCTGGGTTTCGCCACCT	67
Qy	5298	ctgacttgagcgtcgatttttctgatgctcgtcagggggcgagcctatggaaaaacgc	5357
Db	66	CTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGAAAAACGC	7
Qy	5358	cagcaa 5363	
Db	6	CAGCAA 1	
RESULT	3		
AL042026			
LOCUS	AL042026	841 bp	mRNA linear EST 29-FEB-2000
DEFINITION	DKFZp434E111_r1 434 (synonym: htes3) Homo sapiens cDNA clone		
	DKFZp434E111 5', mRNA sequence.		
ACCESSION	AL042026		
VERSION	AL042026.1	GI:5421372	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

REFERENCE	1	(bases 1 to 841)	
AUTHORS	Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.		
TITLE	EST (Poustka, et al.)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Poustka A.J. Department Lehrach Max-Planck-Institute for Molecular Genetics Innestrasse 73, 14195 Berlin, Germany Tel: +49-30-84131623 Fax: +49-30-84131128 Email: poustka@mping-berlin-dahlem.mpg.de This is the 5' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. No sl sequence available. This clone (DKFZp434E111) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.		
FEATURES	Location/Qualifiers		
source	1..841		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="DKFZp434E111"		
	/clone_lib="434 (synonym: htes3)"		
	/tissue_type="testis"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"		
BASE COUNT	226 a 192 c 202 g 221 t		
ORIGIN			
Query Match	14.6%;	Score 841;	DB 9; Length 841;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 841;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	3639	tattgtttattttctaaatacatcaaaatgtatccgctcatgagacaataaccctg	3698
Db	1	TATTTGTTTATTTTCTAAATACATTCAAAATATGTATCCGCTCATGAGACAATAACCCGTG	60
Qy	3699	ataaatgcttcaataatattgaaaagggaagagatgagtatgagtatccaattccgtgtcgc	3758
Db	61	ATAAATGCTTCAATAATATTGAAAAGGAAGAGATGAGTATTCAACATTTCCGTGTGCG	120
Qy	3759	ccttattcccttttttgcggcattttgccttcctgtttttgtcaccaccagaaacgtggt	3818
Db	121	CCTTATTCCCTTTTTCGCGCATTTTGCCCTTCCTGTTTTCCTACCCAGAAACGCTGGT	180
Qy	3819	gaaagtaaaagatgctgaagatcagttgggtgcacgagtggttatacatcgaactggtatct	3878
Db	181	GAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCT	240
Qy	3879	caacagcgggtaagatcccttgagagtttttcgccccgaagaacgttttccaatgatgagcac	3938
Db	241	CAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCCGAAGAACGTTTTCCAATGATGAGCAC	300
Qy	3939	ttttaagttctgtatgtggcgcggtattatcccgattattgacgcggggcaagagcaact	3998
Db	301	TTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGCAAGAGCAACT	360
Qy	3999	cggtcgcgcatacactattctcagaatgacttggtgagtagtactcaccagtcacagaaaa	4058
Db	361	CGGTCGCCGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAAGTCACAGAAA	420
Qy	4059	gcatcttacgggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtga	4118
Db	421	GCATCTTACGGATGGCATGACAGTAGAGAATTATGCAGTGTCTGCCATAACCATGAGTGA	480
Qy	4119	taacactcggcgccaaacttacttctgacaacgatcggaggaccgaaggagtaaccgcttt	4178

QY 4995 agtggctgctgcagtgccgataaagtcgtgtcttaccgggttgactcaagacgatagtt 5054
|||||
Db 774 AGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACC GGTTGGACTCAAGACGATAGTT 715

QY 5055 accggataaggcgcagcgggtcgggctgaacggggggttcgtgcacacagcccgcttggga 5114
|||||
Db 714 ACCGGATAAGGCGCAGCGGTCCGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGGa 655

QY 5115 gcgaacgacctacacccgaactgagataacctacacagcgtgagctatgagaaagcgcacgct 5174
|||||
Db 654 CGCAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCT 595

QY 5175 tcccgaaggagaaaggcggacagggtatccgggtaagcggcaggggtcggaaacaggagagcg 5234
|||||
Db 594 TCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGGAACAGGAGCG 535

QY 5235 cacgaggagctccaggggggaaacgcctggtatctttatagtcctgtcggttttcgcca 5294
|||||
Db 534 CACGAGGGAGCTCCAGGGGGAACGCCTGGTATCTTTATAGTCTCTGCGGTTTCGCCA 475

QY 5295 cctctgactgagcgtcgatttttgtgatgctcgtcagggggcgaggcctatggaaaaa 5354
|||||
Db 474 CCTCTGACTTGAGCGTCGATTTTGTGATGTCGTCAGGGGGCGGAGCCTATGGAAAAA 415

QY 5355 cgccagcaacggcgcccttttacgggttcctcgcccttttgctggcccttttgctcacatggt 5414
|||||
Db 414 CGCCAGCAACGCGCCCTTTTACGGTTCCTGGCCCTTTTGTGCGCCCTTTTGCTCACATGTT 355

QY 5415 ctttccctgcttatccctcgattctgtggataaaccgtattaccgccccttgagtgagctga 5474
|||||
Db 354 CTTTCCCTCGCTTATCCCTCGATTCTGTGGATAACCGTATTACCGCCCTTTTGAGTGAGCTGA 295

QY 5475 taccgctcgcgcagccgaacgacgagcgcagcagtcagtcagtgagcaggaagcgggaaga 5534
|||||
Db 294 TACCGCTCGCGCAGCCGGAACGACGAGCGAGCGAGTCAGTGAGCGGAAGCGGGAAGA 235

QY 5535 gcgcccataacgcaaacccgctctcccgcgcgcttgccggttgccgattcattaatgcag 5588
|||||
Db 234 GCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTTCATTAATGTCAG 181

RESULT 7
AJ281480/c
LOCUS AJ281480 1004 bp mRNA linear EST 30-JUN-2000
DEFINITION 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae CDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION AJ281480
VERSION AJ281480.1 GI:6929360
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE 1 (bases 1 to 1004)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .1004
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"

/clone="4A3A-P4G8"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 252 a 262 c 244 g 244 t 2 others
ORIGIN

Query Match 12.5%; Score 723; DB 9; Length 1004;
Best Local Similarity 99.8%; Pred. NO. 0;
Matches 823; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4568 gtgcctcactgattaagcattggtaactgtcagaccacgaatttactatatactattaga 4627
|||||
Db 825 GTGCCTCACTGATTAAGCATTTGGTAACTGTGACACCACCAAGTTTACTCATATATACTTTAGA 766

QY 4628 ttgatttaaaacttcatttttaatttaaaaggatctagtgaaagatcccttttgataatc 4687
|||||
Db 765 TTGATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATC 706

QY 4688 tcatgacccaaatcccttaaacgtgagtttctgctccactgagcgtcagaccccgtagaaa 4747
|||||
Db 705 TCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCTCCACTGAGCGTCAGACCCCGTAGAAA 646

QY 4748 agatcaaaggatctcttgagatccctttttctgcgcgtaactctgctgcttgcacaaacaa 4807
|||||
Db 645 AGATCAAAGGATCTTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTGCAACAA 586

QY 4808 aaaaaccaccgctaccagcgggtggttcttgccggatcaagagcgtaccacactcttttc 4867
|||||
Db 585 AAAAAACCACCGCTACCAGCGGTGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTC 526

QY 4868 cgaaggtaactggcttcagcagagcgcagataccaaaatactgtccttctagtgtagccgt 4927
|||||
Db 525 CGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTTCTTCTAGTGTAGCCGT 466

QY 4928 agttaggccaccacttcaagaactctgttagcaccgcctacatacctcgtctgttaatcc 4987
|||||
Db 465 AGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCC 406

QY 4988 tgttacagtgctgctgccagtgccgataagtcgttcttacccgggttgactcaagac 5047
|||||
Db 405 TGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTTCTTACCGGGTGGACTCAAGAC 346

QY 5048 gatagttaccgataaaggcgcagcgggtcggctgaacgggggttcgtgcacacagccca 5107
|||||
Db 345 GATAGTTACCGGATAAGGCGCAGCGGTCCGSGCTGAACGGGGGTTCGTGCANACAGCCCA 286

QY 5108 gcttgagcgaacgacctacacccgaactgagataacctacagcgtgagctatgagaaagcg 5167
|||||
Db 285 GCTTGGAGCGAAGCAGCCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCG 226

QY 5168 ccacgcttcgccgaaggagaaaggcggaggtatcccggttaagcggggttcggaacag 5227
|||||
Db 225 CCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTTCGGAACAG 166

QY 5228 gagagcgcagaggagagcttccaggggggaaacgcctggtatctttatagtcctgcgggt 5287
|||||
Db 165 GAGAGCGCACAGGGGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTGTCTGGGT 106

QY 5288 ttcgccacctgtgacttgagcgtcgatttttctgatgctcgtcagggggcgagcctat 5347
|||||
Db 105 TTCGCCACCTCTGACTTGAGCGTCGATTTTGTGTGATGCTCGTCAGGGGGCGGAGCCTAT 46

QY 5348 ggaaaaaacgccagcaacgcggccctttttacgggttcctggttcctttt 5392

```
|||||
Db 45 GGAAGACGCCAGCAACGGCCCTTTTACGGTTCCTGGCCCTTT 1

RESULT 8
BM438846
LOCUS
DEFINITION Iplvr00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
ACCESSION BM438846
VERSION BM438846.1 GI:18460568
KEYWORDS EST.
SOURCE channel catfish.
ORGANISM Ictalurus punctatus
REFERENCE 1 (bases 1 to 1013)
AUTHORS Feng,J., Kucuktas,H., Kocabas,A., Li,P. and Liu,Z.
TITLE Transcriptome of channel catfish (Ictalurus punctatus): initial
JOURNAL analysis of expressed sequence tags from the liver
COMMENT Unpublished (2002)
Contact: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: M13 Reverse.
FEATURES
source
1..1013
/organism="Ictalurus punctatus"
/db_xref="taxon:7998"
/clone_lib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: Noti;
Site_2: Sali"

BASE COUNT 273 a 228 c 245 g 266 t 1 others
ORIGIN

Query Match 12.0%; Score 691; DB 10; Length 1013;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 931; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 3826 aaagatgctgaagatcagtggtggcgacgagtggttacatcgaaactggatctcaacagc 3885
|||||
Db 49 AAGATGCTGAAGATCAGTTGGTGGCAGCAGTGGGTACATCGNACTGGATCTCAACAGC 108

Qy 3886 ggtaagatccttgagagttttcgcccggaagaacggttttccaatgatgagcacttttaa 3945
|||||
Db 109 GGTAAAGATCCTTGAGAGTTTTCGCCCCGAAGAAGCTTTTCCATGATGAGCACTTTTAA 168

Qy 3946 gttctgctatgtggcgcggtattatcccgatttgacgcccgggcaagagcaactcggtcgc 4005
|||||
Db 169 GTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCCG 228

Qy 4006 cgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaaagcatctt 4065
|||||
Db 229 CGCATACACTATTCTCAGAAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAAGCATCTT 288

Qy 4066 acggatggcatgacagtaagagaattatcgagtgctgccataaccatgagtgataacact 4125
|||||
Db 289 ACGGATGCGATGACAGTAAGAGAAATATATGTCAGTGTCTGCCATAACCATGAGTGATAACT 348

Qy 4126 gcggccaacttacttctgacaaacgatcgggagaccgaaggagcgttaaccgctttttgca 4185
|||||
Db 349 GCGGCCCACTTACTTCTGACAAACGATCGGAGGACCGAA-GAGCTAACCGCTTTTTCGCAC 407

Qy 4186 aacatgggggatc-atgtaactcgcttgatcggttgggaacccgagctgaatgaagccat 4244
|||||
```

```
Db 408 AACATGGGGATCAATGTAACCTGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCAT 467

Qy 4245 accaaacgacgagcgtgacacaccacgacgctgtagcaatggcaacaacgttgcgcaact 4304
|||||
Db 468 ACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCCGAACT 527

Qy 4305 attaaatggcgaactacttactactctagcttccccgggcaacaattaatagactggatggaggc 4364
|||||
Db 528 ATTAACCTGCGGAACACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGC 587

Qy 4365 ggataaaagtgcaggaccacttctcgctcgcccttcccgctggtggtggtttattgtctga 4424
|||||
Db 588 GGATAAAGTTCAGGACCACCTTCTGCGCTCGGCCCTTCCGGCTGGCTGTTATTGCTGA 647

Qy 4425 taaatctgagccggtgagcgtgggtctcgcggttatcattgcagcactgggcccagatgg 4484
|||||
Db 648 TAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGG 707

Qy 4485 taagccctccgctatcgtagttatctcacgacgggggagtcaggccaactatggatgaacg 4544
|||||
Db 708 TAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGAGTTCAGGCAACTATGGATGAACG 767

Qy 4545 aaatagacagatcgctgagataggtgctcactgattaaagcattggttaactgtcagacca 4604
|||||
Db 768 AAATAGACAGATCGCTGAGATAGTGGTGCCTCACTGATTAAAGCATTTGGTAAGTGTGACAGCA 827

Qy 4605 agtttactcatatatacttttagattgatttaaaacttcatcttttaatttaaaaggatcta 4664
|||||
Db 828 AGTTTACTCATATATACTTTTAGATTGATTAAAACTTCATTTTAAATTAAGGATCTA 887

Qy 4665 ggtgaagatccttttggataatctcatgacccaaaaccccttaacgtgagtttctgtcca 4724
|||||
Db 888 GGTGAAGATCCTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTCGTTCCA 947

Qy 4725 ctgagcgtcagacccccgtagaaaagatcaaaagg 4757
|||||
Db 948 CTGAGCGTCAGACCCCGTAGAAAAGATCAAAAGG 980

RESULT 9
AU081040/c
LOCUS
DEFINITION AU081040 Oncorhynchus mykiss Kidney infected by infectious
hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KA9,
mRNA sequence.
ACCESSION AU081040
VERSION AU081040.1 GI:6431388
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1039)
AUTHORS Kono,T., Sakai,M. and LaPatra,S.E.
TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious
Hematopoietic Necrosis Virus
JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES
source
1..1039
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KA9"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"
```


BASE COUNT	258 a	256 c	260 g	265 t	
ORIGIN					
Query Match	12.0%; Score 691; DB 9; Length 1039;				
Best Local Similarity	100.0%; Pred. NO. 0;				
Matches 691; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY 5077	ggctgaacgggggttcgtgcacacagcccagcttgagcgaaacgacctacacccaactg 5136				
Db					
Db 1039	GGCTGAACGGGGGTTCGTGCACACAGCCCGCTTGGAGCGAACGACCTACACCGAACTG 980				
QY 5137	agatacctacagcgtgagctatgagaaaagcgccacgcttccccgaaggagaaagcgagac 5196				
Db					
Db 979	AGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGCGAC 920				
QY 5197	aggtatccggtaagcggcaggggtcggaacagagagcgacgagggagcttccaggggga 5256				
Db					
Db 919	AGGTATCCGGTAAGCGGCAGGGTCGGAACACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA 860				
QY 5257	aacgcctggtatctttatagtcctgtcggggtttcgccacctctgacttgagcgtcgattt 5316				
Db					
Db 859	AAGCCTGGTATCTTTATAGTCTCTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTT 800				
QY 5317	tttgtatgctcgtcagggggcgagcctatggaaaaacgccagcaacgcgcccttttta 5376				
Db					
Db 799	TTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGAANAACGCCAGCAACGCGGCCCTTTTA 740				
QY 5377	cggttcctggccttttgctggccttttgctcacatgttcttcctgcgttatccccctgat 5436				
Db					
Db 739	CGGTTCCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTCTCGGTTATCCCCCTGAT 680				
QY 5437	tctgtggataaacgctattaccgcctttgagtgagctgataccgcctcgcgcagccgaacg 5496				
Db					
Db 679	TCTGTGGATAACCGTATTACCGCCTTTGAGTGAGTGAGTGATACCGCTCGCCGCGACCGGAACG 620				
QY 5497	accgagcgcagcagtcagtcagtcagcaggaagcgggaagcgcgcccaatacgcacaaacgcct 5556				
Db					
Db 619	ACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCT 560				
QY 5557	ctccccgcgcttgccgattcattaatgcagctggcacgacaggttccccgactggaaa 5616				
Db					
Db 559	CTCCCCGCGCTTGCCCGATTTCATTAAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAA 500				
QY 5617	gcgggcagtcagcgaacgcaattaatgtgagttagctcactcattaggcaccgccagct 5676				
Db					
Db 499	CGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCACTATTAGGCACCCCGAGCT 440				
QY 5677	ttacactttatgcttcgggctcgtatgttgtgtggaattgtgagcgggataacaatttcac 5736				
Db					
Db 439	TTACACTTTATGCTTCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTTAC 380				
QY 5737	acaggaacacgctatgaccatgattacgcca 5767				
Db					
Db 379	ACAGGAACACAGCTATGACCATGATTACGCCA 349				
RESULT 10					
BI753192					
LOCUS	BI753192	780 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION	603026066F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196256 5', mRNA sequence.				
ACCESSION	BI753192				
VERSION	BI753192.1 GI:15744770				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 780)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				

COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11491 row: d column: 17 High quality sequence stop: 780.				
FEATURES	Location/Qualifiers				
source	1. .780				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:5196256"				
	/clone_lib="NIH_MGC_114"				
	/lab_host="DH10B"				
	/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."				
BASE COUNT	111 a	255 c	260 g	154 t	
ORIGIN					
Query Match	11.9%; Score 686; DB 10; Length 780;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY 844	ccccgcagtgctgctgcacacgcttccctccctggcttccccactccttctcctcctc 903				
Db					
Db 54	CCCCCGCAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCACACTCCTTCTCCTCCTC 113				
QY 904	tggctcctgggtgagagagtggggctgagggcgaggagagatgcagagctgctggtgacg 963				
Db					
Db 114	TGGCTCCTGGGTGGAGGAGTGGGGGTGAGGGCGGGGAGGATGCAGAGCTGCTGGTGACG 173				
QY 964	gtcgtggggcgcgctgcgggcattcgcctgaagacccccgggggcccctgctctgct 1023				
Db					
Db 174	GTGCGTGGGGCGGCTGCGGGCATTCGCCTGAAGACCCCGGGGGCCCTGCTCTGCT 233				
QY 1024	ttcctgggcatcccccttgcggagccaccatgggacccccgccttcttctgcccaggag 1083				
Db					
Db 234	TTCTTGCGCATCCCTTTGCGGAGCCACCCCATGGACCCCGCTGCTTCTGCCACCGGAG 293				
QY 1084	cccaagcagccttggtcagggggtggtagacgcgtacacaccttccagagtgctgctaccaa 1143				
Db					
Db 294	CCCAAGCAGCCTTGGTCAGGGGTGGTAGACCGCTACAACCTTCCAGAGTGTCTGTACCAA 353				
QY 1144	tatgtggacacccctataccccagggttttgaggcaccgagatgtggaaccccaaccgtgag 1203				
Db					
Db 354	TATGTGGACACCTATACCCAGGTTTGTAGGGCACCGGAGATGTGGAACCCCAACCGTGAG 413				
QY 1204	ctgagcaggactgcctgtacctcaacgtgtggacacaccatacccccgccctacatcccc 1263				
Db					
Db 414	CTGAGCGAGGACTGCCTGTACCTCAACGCTGTGGACACCATACCCCGGCGCTACATCCCC 473				
QY 1264	acccctgctcctgcttgatctatgggggttggttcttacagtggggcccctcctcttgac 1323				
Db					
Db 474	ACCCCTGTCCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTTCTCCTTGGAC 533				
QY 1324	gtgtacgatggccgcttcttgggtacagggcgagagagactgtgctggtgtccatgaactac 1383				
Db					
Db 534	GTGTACGATGGCGGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATGAACACTAC 593				
QY 1384	cgggtggagcccttggcttccctgcccctgcccgggagccgagagccccggggcaatgtg 1443				

Db 594 CCGGTGGAGCCTTTGGCTTCCTGGCCCTGCCGGGAGCGGAGAGCCCGGGCAATCTG 653

QY 1444 ggtctcctggatcagagctggccctgcagtggtgcaggagaacgtggcagcccttcggg 1503

Db 654 GGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCCTTCGGG 713

QY 1504 ggtgacccgacatcagtgacgtgtt 1529

Db 714 GGTGACCCGACATCAGTGACGCTGTT 739

RESULT 11

AU081044/c

LOCUS AU081044 1163 bp mRNA linear EST 30-AUG-2001

DEFINITION AU081044 Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KBL, mRNA sequence.

ACCESSION AU081044

VERSION AU081044.1 GI:6431392

KEYWORDS EST.

SOURCE rainbow trout.

ORGANISM Oncorhynchus mykiss

REFERENCE 1 (bases 1 to 1163)

AUTHORS Kono,T., Sakai,M. and LaPatra,S.E.

TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus

JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)

COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakalecc.miyazaki-u.ac.jp.

FEATURES

source

1. .1163

/organism="Oncorhynchus mykiss"

/db_xref="taxon:8022"

/clone="KBL"

/clone_lib="Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus"

/tissue_type="Kidney infected by infectious hematopoietic necrosis virus"

BASE COUNT 272 a 307 c 298 g 286 t

ORIGIN

Query Match 11.8%; Score 682; DB 9; Length 1163;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 682; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5086 gggggttcgtgcacacagcccgcttgagcgaacgacctacacggaactgagataccta 5145

Db 1050 GGGGTTTCGTGCACACAGCCCGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTA 991

QY 5146 cagcgtgagctatgagaaagcgccacgcttcccgaaggagaaagcgaggtatccg 5205

Db 990 CAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCG 931

QY 5206 gtaagcggcagggtcggaaacagagagcgacgagggagcttccagggggaaacgcctgg 5265

Db 930 GTAAGCGCAGGGTCGGACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAACGCCCTGG 871

QY 5266 tatctttatagtcctgcgggttcgccacctctgacctgagcgtcgattttgtgatgc 5325

Db 870 TATCTTTATAGTCCCTGTCGGGTTTCGCCACCTCTGACTTGACCGTCGATTTTGTGATGC 811

QY 5326 tcgtcaggggggcggagcctatggaaaaacgccagcaacggcgccctttttacggttcctg 5385

Db 810 TCGTCAGGGGGCGGAGCCCTATGGAANAACGCCAGCAACGGCGCCTTTTACGGTTCCCTG 751

QY 5386 gccttttgctggccttttgcacatgttcttccctgcgttatccctgatctctgtgat 5445

Db 750 GCCTTTTGCTGGCCTTTTGCTCACATGTTCTTCTTCCGCTTATCCCTGATTCTGTGGAT 691

QY 5446 aaccgtattaccgcctttgagtgagctgataccgcctgcgcgagccgaacgacgagcgc 5505

Db 690 AACCGTATTACCGCCTTTGAGTGAGTGAGTGATACCGCTCGCCGCGAGCGGAACGACCGAGCGC 631

QY 5506 agcagtcagtgagcgaaggaagcgaagagcgcccaatacgcgaacccgcctctccccgcg 5565

Db 630 AGCGAGTCAGTGAGCGAGGAGCGGAAGAGAGCGCCCAATACGCAACCGCCTCTCCCCGCG 571

QY 5566 cgttgccgattcattaatgcagctggcacgacaggttccccgactggaaaagcgggcagt 5625

Db 570 CGTTGGCGGATTCAATATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGCAGT 511

QY 5626 gagecgaacgcaattaatgtgagttagctcactcattaggcaccgccaggttttacacttt 5685

Db 510 GAGCGCAACGCAATTAAATGTGAGTTAGCTCACTCATTAGGCACCCCGAGGCTTTACACTTT 451

QY 5686 atgcttccggctcgtatgttgtgtgaattgtgagcggataacaatttcacacaggaaac 5745

Db 450 ATGCTTCCGGCTCGTATGTTGTGTGGAATGTGACGGGATAACAATTTCACACAGGAAAC 391

QY 5746 agctatgaccatgattacgcca 5767

Db 390 AGCTATGACCATGATTACGCCA 369

RESULT 12

AL646751/c

LOCUS AL646751 XGC-neurula Silurana tropicalis cDNA clone TNeu029j21 5', mRNA sequence.

DEFINITION AL646751 680 bp mRNA linear EST 13-DEC-2001

ACCESSION AL646751

VERSION AL646751.1 GI:17654643

KEYWORDS EST.

SOURCE western clawed frog.

ORGANISM Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Silurana.

REFERENCE 1 (bases 1 to 680)

AUTHORS Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.

TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Huckle E
Sanger Centre
Hinxtton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu029j21.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.

FEATURES

Location/Qualifiers

1. .680

/organism="Silurana tropicalis"

/db_xref="taxon:8364"

/clone="TNeu029j21"

/clone_lib="XGC-neurula"

/dev_stage="neurula"

/lab_host="Escherichia coli DH10B"

/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

BASE COUNT 144 a 201 c 189 g 146 t

ORIGIN

Query Match 11.8%; Score 679; DB 9; Length 680;

Best Local Similarity 100.0%; Pred. No. 0;

Matches	679;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	4940	acttcaagaactctgtagcaccgcctacatacctcgctctgtctaactctgttaccagtg	4999						
Db	680	ACTTCRAAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCTGTTACCACTGG	621						
QY	5000	ctgctgccagtggcgataaagtcgtgttcttaccgggttggaactcaagacgatatgttacc	5059						
Db	620	CTGCTGCCAGTGGCGATAAGTCGTGCTTACC	561						
QY	5060	ataagggcagcggctcgggtgaacgggggttcgtgcacacagccagcttggaagcga	5119						
Db	560	ATAAGGCGCAGCGGTTCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAA	501						
QY	5120	cgacctacaccgaactgagatacctacagcgtgagctatgagaagcgcacgcttcccg	5179						
Db	500	CGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGNAAGCGCCACGCTTCCCG	441						
QY	5180	aaggggaaagggcggacaggtatccgggtaagcggcaggggtcggaacaggaagcgcacga	5239						
Db	440	AAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGACGGTTCGGAACAGGAGCGCACGA	381						
QY	5240	gggagcttccaggggggaaacgcctgggtatctttatagtcctgtcggttccgcaacctct	5299						
Db	380	GGGAGCTTCCAGGGGAAACGCCCTGATCTTATAGTCCTGTGCGGTTTCGCCACCTCT	321						
QY	5300	gacttgagcgtcgatttttgtgatgctcgtcagtcagggggcgagcctatggaacacgcga	5359						
Db	320	GACTTGAGCGTCGATTTTGTGATGTCGTACAGGGGGGGAGGCTATGGAAAAACGCCA	261						
QY	5360	gcaacggcgccttttttacgggttcctggcccttttgcggccttttgcctcacatgttcttc	5419						
Db	260	GCAACGGCGCCCTTTTACGGTTCCTGGCCCTTTTGTGGCCCTTTGCTCACATGTTCTTTC	201						
QY	5420	ctgcgttatccctgattctgtggaataaccgtattaccgcctttgagtgagctgataccg	5479						
Db	200	CTGCGTATCCCTCTGATTCCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCG	141						
QY	5480	ctcgcgcagccgaaacgacgagcgcagtcagtcagtcaggaagcgaagcgcgc	5539						
Db	140	CTCGCCGAGCCGGAACGACCGAGCGCAGCGAGTCAGTGAGCGGAGGAACGGAAGCGCC	81						
QY	5540	caatacgcaaacccgcctctccccgcgcttgcccgattcattaatgcaagtcggaagcaga	5599						
Db	80	CAATACGCAAAACCGCCTCTCCCCCGCGGTGGCCGATTTCATTATGTCAGCTGGCAGACA	21						
QY	5600	ggtttcccgactggaaagc	5618						
Db	20	GGTTTCCCGACTGGAAAGC	2						

RESULT 13
AL656688/c
LOCUS
DEFINITION AL656688 XGC-neurula Silurana tropicalis cDNA clone TNeu032h21 5', mRNA sequence. 707 bp mRNA linear EST 13-DEC-2001
ACCESSION AL656688
VERSION AL656688.1 GI:17669128
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 707)
AUTHORS Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: TNeu032h21.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1. .707
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu032h21"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 153 a 205 c 195 g 154 t
ORIGIN

Query Match	11.8%;	Score	678;	DB	9;	Length	707;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	678;	Conservative	0;	Mismatches	0;	Indels	0;
QY	4912	cttctagtgtagccgttagtgagccaccacttcaagaactctgtagcaccgcctacatac	4971				
Db	678	CTTCTAGTGTAGCCGTAGTTAGGCCACCACCTCAAGAACTCTGTAGCACCGCCTACATAC	619				
QY	4972	ctcgcctctgtaatacctgttaccagtggtcgtcgtccagtggcgataaagtcgtgtcttacc	5031				
Db	618	CTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTTTACC	559				
QY	5032	gggttggaactcaagacgacgatagttaccggataaaggcgcagcggctggaacgggggt	5091				
Db	558	GGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGT	499				
QY	5092	tcgtgcacacagcccagcttgagcgaaagacccctacacccaactgagataccctacagcgt	5151				
Db	498	TCGTGCACACAGCCCCAGCTTGGAGCGGAACGACCTACACCGAACTGAGATACCTACAGCGT	439				
QY	5152	gagctatgagaaagcgcacgccttcccgaaggagaaaggcgcaggtatccggtaagc	5211				
Db	438	GAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGGGACAGGTATCCGGTAAGC	379				
QY	5212	ggcagggtcggaacagagagcgcagcagggagcttccagggggaacgcctggtatctt	5271				
Db	378	GGCAGGGTTCGGAACAGGAGAGCGCACGAGGAGGCTTCCAGGGGGAACGCGCTGGTATCTT	319				
QY	5272	tatagtcctgctgggttttcgccacctctgaccttgagctgagctgattttgtgatgctcgta	5331				
Db	318	TATAGTCTCTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCA	259				
QY	5332	ggggggcggagcctatggaaaaacgcgcgaacgcggcctttttacggttccctggccttt	5391				
Db	258	GGGGGGCGGAGCCTATGGAAAAACGCCACGCAACGCGGCTTTTACGGTTCTCTGGCCTTT	199				
QY	5392	tgctggccttttgctcacatgttcttcttcctgcgttatccccctgattctgtggataaccgt	5451				
Db	198	TGCTGGCCTTTTGCTCACATGTTCTTTTCTCGGTTATCCCCCTGATTCTGTGGATAACCGT	139				
QY	5452	attaccgcctttgagtgagctgatacgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgag	5511				
Db	138	ATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCGCAGCCGAAACGACGCGCAGCGAG	79				
QY	5512	tcagtgagcgaagcgggaagcgcgccaatacgcgaacccgccttccccgcgcttgg	5571				
Db	78	TCAGTGAGCGAGGAAGCGGAAGAGCGGCCCAATACGAAACCGCCTCTCCCCCGCGCTTGG	19				
QY	5572	ccgattcattaatgcagc	5589				
Db	18	CCGATTCAATTAATGCAGC	1				

```
RESULT 14
AJ2811449/c
LOCUS
DEFINITION 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION AJ281149
VERSION AJ281149.1 GI:6929329
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE 1 (bases 1 to 800)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoerge,W., Soares,M.B.
and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.
FEATURES
source
1. .800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4D5"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 203 a 206 c 198 g 193 t
ORIGIN
Query Match 11.8%; Score 678; DB 9; Length 800;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 778; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4620 actttagattgatttaaaacttcatttttaatttaaaggatctaggtgaagatcctttt 4679
|||||
Db 800 ACTTTAGATTGATTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 741
QY 4680 tgataatctcatgaccaaatacccttaacgtgagtttcttcactgagcgtaagacc 4739
|||||
Db 740 TGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACC 681
QY 4740 cgtagaaaagatcaaaaggatctcttgagatccctttttctgcgcgtaatctgctgt 4799
|||||
Db 680 CGTAGAAAAGATCAAAGGATCTCTTGAGATCCTTTTCTGCGGTAATCTGCTGCTT 621
QY 4800 gcaacaaaaaaaccaccgctaccagcggtgtgtttgttcgggatcaagagctaccaac 4859
|||||
Db 620 GCAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAAC 561
QY 4860 tctttttccgaaggtaactggcttcagcagagcgagataccaaatactgtccttctagt 4919
|||||
Db 560 TCTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGT 501
QY 4920 gtagccgtagttaggccaccacttcaagaactctgtagcagcgccctacatacctcgtct 4979
|||||
```

```
Db 500 GTAGCCGTAGTTAGCCACCACCTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCT 441
QY 4980 gctaactcctgttaccagtggtcgtcgtgccagtggcgataagtcgtgtcttaccgggttga 5039
|||||
Db 440 GCTAATCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTTTACCGGGTTGA 381
QY 5040 ctcaagacgatagtttaccgggataaaggcgagcggctcgggctgaacggggggttcgtgcac 5099
|||||
Db 380 CTCAAGACGATAGTTACCGGATAAGGCGCAGGGTCGGGGTGAACGGGGGGTTCGTGCAC 321
QY 5100 acagccagcttgagcgaaacacacacacacacacacacacacacacacacacacacacacac 5159
|||||
Db 320 ACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATG 261
QY 5160 agaagcgccacgcttcccgaagggagaaaagcgaggtatccggttaagcggcaggggt 5219
|||||
Db 260 AGAAAGCGCACACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGGGT 201
QY 5220 cggaacagagagcgacgagggagcttccaggggggaaacgcctgtatctttatagtc 5279
|||||
Db 200 CGGAACAGGAGAGCGCAGGAGGAGCTTCCAGGGGAAACGCCTGTATCTTTATAGTCC 141
QY 5280 tgtcgggttccgcccactctgactgagcgtcgatttttgtgatgctcgtcagggggcg 5339
|||||
Db 140 TGTCGGGTTTCGCCACCTCTGACTTGACGCTCTATTTTGTGATGCTCGTCAGGGGGCG 81
QY 5340 gagcctatggaaaaacgccagcaacgcggccttttttacggttcctgccttttgcctgccc 5399
|||||
Db 80 GAGCCTATGAAAAAACGCCAGCAACGCGGCTTTTACGGTTCTCTGGCCTTTTGCTGGCC 21
RESULT 15
AL645114/c
LOCUS
DEFINITION AL645114 XGC-neurula Silurana tropicalis cDNA clone TNeu017b19 5',
mRNA sequence.
ACCESSION AL645114
VERSION AL645114.1 GI:16797239
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 724)
AUTHORS Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxtton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu017b19.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1. .724
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu017b19"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 162 a 206 c 194 g 162 t
ORIGIN
```

Query Match		11.7%;	Score 673;	DB 9;	Length 724;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 723;	Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;	
QY	4800	gcaaaacaaaaaacaccgctaccagcgggtggtttgtttgccggatcaagagctaccaac				4859
Db	724	GCAAAACAAAAAACACCACCGCTACCAGCGGTGGTTGTTGCCGGATCAAGAGCTACCAAC				665
QY	4860	tcttttccgaaggttaactggcttcagcagagcgcagataccaaataactgtccttctagt				4919
Db	664	TCTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGT				605
QY	4920	gtagccgtagttagccaccacttcagaactctgtagcaccgcctacatacctcgctct				4979
Db	604	GTAGCCGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCT				545
QY	4980	gctaatacctgttacagtggtgctgctgccagtggcgataaagtcgtcttaccgggttggga				5039
Db	544	GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGA				485
QY	5040	ctcaagacgatatgttaccggataaagcgcagcgggtcgggctgaacgggggttcgtgcac				5099
Db	484	CTCAAGACGATAGTTACCGGATAAGCGCGCAGCGTCCGGCTGAACGGGGGGTTCGTGCAC				425
QY	5100	acagcccagcttgagcgaacgacacctacaccgaaactgagataacctacagcgtgagctatg				5159
Db	424	ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATG				365
QY	5160	agaaagccgcacgcttccccgaagggaagaaagcggaaggtatccggtaagcggcaggggt				5219
Db	364	AGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGT				305
QY	5220	cggaacagggagcgcacgagggagcttccaggggggaaacgccttggtatctttatagtc				5279
Db	304	CGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCC				245
QY	5280	tgtcgggtttcgcacacctctgacttgagcgtcgatttttgtgatgctcgtcagggggcg				5339
Db	244	TGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGTGATGCTCGTCAGGGGGCG				185
QY	5340	gagcctatggaaaaaacgccagcaacgcggcctttttacggttcctggccttttgcgtgcc				5399
Db	184	GAGCCTATGGAAAAAACGCCAGCAACGGGCCTTTTACGGTTCTCTGGCCTTTTGTGGCC				125
QY	5400	ttttgctcacatgttcttctcctgcgttatccccctgattctgtggataaccgtattaccgc				5459
Db	124	TTTTTGCTCACATGTTCTTCTCCTGCGTTATCCCTGATTCTGTGGATAACCCGTATTACCGC				65
QY	5460	ctttgagtgagctgataccgctcgcgcagccgaacccgagcgcagtcagtcagtcag				5519
Db	64	CTTTGAGTGAGCTGATACCGCTCGCCGCGACCGGAACCGACCGGAGCGGTCAGTGAG				5
QY	5520	cgag 5523				
Db	4	CGAG 1				

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 22:37:24 ; Search time 154.73 Seconds
(without alignments)
9155.099 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcatgcctgcaggtc.....ctatgaccatgattacgcca 5767

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2637	45.7	8854	3	US-09-053-549-1
2	2630	45.6	6295	2	US-08-659-206A-4
3	2628	45.6	4950	3	US-08-789-333F-58
4	2628	45.6	9639	4	US-09-147-208-26
5	2533	43.9	4283	1	US-08-343-401A-3
6	2533	43.9	4283	1	US-08-445-265A-1
7	2533	43.9	4283	3	US-08-990-442-1
8	2527	43.8	7566	2	US-08-232-016-23
9	2527	43.8	7639	2	US-08-232-016-22
10	2526	43.8	4713	4	US-09-194-285-7
11	2526	43.8	4724	4	US-09-194-285-8
12	2482	43.0	7387	4	US-09-238-356-28
13	2480	43.0	4045	4	US-08-464-700-54
14	2475	42.9	7560	4	US-08-844-274-20
15	2448	42.4	11284	3	US-08-978-741-5
16	2418	41.9	8225	4	US-08-793-618-1
17	2413	41.8	6350	2	US-08-385-335A-8
18	2398	41.6	3819	4	US-09-042-353-393
19	2398	41.6	3819	4	US-08-758-417A-243
20	2398	41.6	3881	4	US-09-042-353-369
21	2398	41.6	3881	4	US-08-758-417A-217
22	2398	41.6	6387	1	US-07-721-775A-1
23	2398	41.6	6387	1	US-08-339-658-1
24	2367	41.0	19307	3	US-08-836-022A-10
25	2367	41.0	19307	4	US-09-427-048A-10
26	2357	40.9	4696	2	US-08-929-967-15
27	2357	40.9	5158	2	US-08-929-967-16

28	2357	40.9	7607	1	US-08-222-616-19	Sequence 19, Appl
29	2357	40.9	7607	5	PCT-US95-04228-19	Sequence 19, Appl
30	2357	40.9	9108	5	PCT-US95-04228-45	Sequence 45, Appl
c 31	2350	40.7	5639	3	US-09-175-690A-1	Sequence 1, Appl
c 32	2350	40.7	6306	5	PCT-US94-00658-1	Sequence 1, Appl
33	2347	40.7	3343	6	5453363-2	Patent No. 5453363
34	2299	39.9	4054	3	US-09-098-287A-9	Sequence 9, Appl
c 35	2248	39.0	4622	4	US-08-509-024-6	Sequence 6, Appl
c 36	2248	39.0	4622	4	US-09-333-279-6	Sequence 6, Appl
c 37	2248	39.0	9641	2	US-08-374-483-3	Sequence 3, Appl
38	2242	38.9	11958	4	US-08-927-317-7	Sequence 7, Appl
39	2240	38.8	8299	1	US-08-462-014-2	Sequence 2, Appl
40	2240	38.8	8299	3	US-08-923-137-3	Sequence 3, Appl
41	2240	38.8	8299	4	US-08-973-334-5	Sequence 5, Appl
42	2240	38.8	8299	4	US-09-563-869A-5	Sequence 5, Appl
43	2240	38.8	8509	1	US-08-462-014-1	Sequence 1, Appl
44	2240	38.8	8509	4	US-08-973-334-4	Sequence 4, Appl
45	2240	38.8	8509	4	US-09-563-869A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-053-549-1
; Sequence 1, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..3694
; OTHER INFORMATION: /product= "hyFLIB protein"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8854
; OTHER INFORMATION: /note= "Sequence of pCIB5520
; Patent No. 6121521
; OTHER INFORMATION: containing coding sequence for hyFLIB protein"

US-09-053-549-1

Query Match		45.7%;	Score 2637;	DB 3;	Length 8854;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2637;	Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;
QY	3131	cgaattcactggccgctcgtttttacaaacgctgactgggaaaaacccctggcgttacccaac	3190		
Db	3892	CGAATTCACTGGCCGTCGTTTACAACTGCTGACTGGGAAACCCCTGGCGTTACCCAAC	3951		
QY	3191	ttaatgccttcagcacacatcccccttcgccagctggcgtaataagcgaagagccccgca	3250		
Db	3952	TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCAAGAGGCCCGCA	4011		
QY	3251	ccgatcgccctcccaacagttgcgaagcctgaatggcggaatggcgctgatcggtatt	3310		
Db	4012	CCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGCGCTGATCGGTATT	4071		
QY	3311	ttctccttaagcatctgtgcggtatttcacacgcgcatatggtgcactctcagtacaatct	3370		
Db	4072	TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACCTCTCAGTACAATCT	4131		
QY	3371	gctctgatgcgcgatagttaagccagcccgacacccgcgaacaccccgctgacgcgccct	3430		
Db	4132	GCTCTGATGCGGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACCGCCCT	4191		
QY	3431	gacgggctgtctgctccccggcatccgcttacagacaaagctgtgacgctctccgggagct	3490		
Db	4192	GACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAAGCTGTGACCGTCTCCGGGAGCT	4251		
QY	3491	gcattgttcagagaggttttcacccgtcatcccggaacccggaacccgcgagacgaagggcctcgtga	3550		
Db	4252	GCATGTGTCAGAGGTTTTCACCGTCTACCCGAAACCGCGAGACGAAAGGGCCCTCGTGA	4311		
QY	3551	tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcagtgcca	3610		
Db	4312	TACGCCTATTTTATAGGTTAATGTATGATATAATAATGGTTTCTTAGACGTCAGGTGGCA	4371		
QY	3611	cttttcggggaatgtgcgcggaacccctatttgtttatttttcaaatattgaaaggaaga	3670		
Db	4372	CTTTTCGGGAAATGTGCGCGGAACCCCTATTGTGTTATTATTTTCTAAATACATTCAAATA	4431		
QY	3671	tgtatccgctcatgagacaataaaccttgataaaatgcttcaataataattgaaaaaggaaga	3730		
Db	4432	TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAGGAAGA	4491		
QY	3731	gtatgagtattcaacatttcgcgtgtcgcccttatcccttttttttggcgcattttgccctc	3790		
Db	4492	GTATGAGTATTCAACATTTCCGCTGTGCGCCCTATTCCCTTTTTTTTGGCGCATTTTGCCCTC	4551		
QY	3791	ctgttttttgctcaccacgaaacgctggtgaaagtataaagatgctgaagatcagttgggtg	3850		
Db	4552	CTGTTTTCGCTCACCCAGAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTG	4611		
QY	3851	cacgagtggtttacatcgaaactggatctcaacacgcggtgaagatccttgagagttttcgcc	3910		
Db	4612	CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC	4671		
QY	3911	ccgaagaaggtttccaatgatgagcaacttttaagttctgctatgtggcgcggtattat	3970		
Db	4672	CCGAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT	4731		
QY	3971	cccgtattgacgcgcgggcaagagcaactcggtcgccgcatacacattctcagaatgact	4030		
Db	4732	CCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAAATGACT	4791		
QY	4031	tggttgagtactcaccagtcacagaaaaagcatcttcacggatggcgtaagagaaat	4090		
Db	4792	TGGTTGAGTACTCACCGAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT	4851		
QY	4091	tatgcagtgtgccataaaccatgagtataaacactcgggccaaacttacttctgacaacga	4150		
Db	4852	TATGCAGTGTGCCATAAACCATGAGTGATAAACACTGCGGCCCAACTTACTTCTGACAAACGA	4911		

QY	4151	tcgaggaccgaaggagctaaaccgcttttttgcaacaacatgggggatcatgtaactcgcc	4210
Db	4912	TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAACATGGGGATCATGTAACCTGCC	4971
QY	4211	ttgatcgttgggaaccggagctgaatgaagccataaccaacgacgagcgtgacaccaga	4270
Db	4972	TTGATCGTTGGGAACCGGAGCTGAATGAAGCATACCAACGACGAGCGTGACACCAGA	5031
QY	4271	tgcctgtagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactag	4330
Db	5032	TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTAG	5091
QY	4331	cttccccggcaacaattaatagactggatggaggcgggataaaagtgcaggaccacttctgc	4390
Db	5092	CTTCCCGGCAACAATTAACTAGACTGGATGGAGGGGGGATAAAGTTGCAGGACCACCTCTGC	5151
QY	4391	gctcgcccttcggcgtggctggttttatgtcgtgataaaactctggagccggtgagcgtgggt	4450
Db	5152	GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCCGCTGAGCGTGGGT	5211
QY	4451	ctcgcggtatcattgcagcactggggccagatggtaagccctcccgatatcgttagttatct	4510
Db	5212	CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT	5271
QY	4511	acacgacgggagtcaggcaactatggatgaacgaataagacagatcgtcgagataggtg	4570
Db	5272	ACACGACGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTG	5331
QY	4571	cctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattg	4630
Db	5332	CCTCACTGATTAAAGCATTTGGTAACGTGTAGACCAAGTTTACTCATATATACTTTAGATTG	5391
QY	4631	atttaaaacttcattttttaatttaaaggatctaggtgaagatccttttttgataaatctca	4690
Db	5392	ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGATAAATCTCA	5451
QY	4691	tgaccaaaaaatcccttaacgtgagtttcttcgcgtgtaactcgtcgttgcgaacaaaa	4750
Db	5452	TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA	5511
QY	4751	tcaaaggatcttcttgagatccctttttcttcgcgtgtaactcgtcgttgcgaacaaaa	4810
Db	5512	TCAAAGGATCTCTTTGAGATCCCTTTTTCGCGCGTAATCTGCTGCTGCAAAACAAAA	5571
QY	4811	aaccacccgtaccagcgggtggtttgtttgcgggatcaagagctaceaaactcttttccga	4870
Db	5572	AACCACCCGTACCAGCGGTGGTTGTTTGCGCGGATCAAGAGCTACCAACTCTTTTCCGA	5631
QY	4871	aggtaaactggcttcagcagcgcagatataccaaaatactgtccttctagtgtagccgtagt	4930
Db	5632	AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCCGTAGT	5691
QY	4931	taggcacaccacttcaaagaactctgtagcaccgcctacatacctcgtctgtctaactcctgt	4990
Db	5692	TAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCCTACATACTCGCTCTGCTAATCCTGT	5751
QY	4991	taccagtggctgctgcccagtgccgataaagtctgtcttaccgggttggaactcaagacgat	5050
Db	5752	TACCACTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACC GG GTTGACTCAAGACGAT	5811
QY	5051	agttaccggataaaggcgcgagcggctcgggctgaaacgggggttcgtgcacacagcccagct	5110
Db	5812	AGTTACCGGATAAGGCGCAGCGGTGCGGCTGGAACGGGGGTTCGTGCACACAGCCCAGCT	5871
QY	5111	tggagcgaacgacctacacccaactgagataacctacagcgtgagctatgagaaagcgcca	5170
Db	5872	TGGAGCGAACGACCTACACCGAACTGAGTACCTACAGCGTGAGCTATGAGAAAGCCCA	5931
QY	5171	cgcttccccgaaggagaaaggcgagcaggtatcccggttaagcggcaggggtcggaaacaggag	5230
Db	5932	CGCTTCCCGAAGGGAGAAAGGGGACAGGTTATCCGGTAAAGCGCAGGGTCGGAACAGGAG	5991

QY 3978 tgacgccgggcaagagcaactcggctgcgccgcatatactattctcagaatgaacttggtga 4037
Db 5455 TGACGCCGGCAAGAGCAACTCGGTCCGCCATACACTATTCTCAGAAAGACTGGTTGA 5396

QY 4038 gtactcaccagtcacagaaaaagcatcttaccggatggcatgacagtaagagaaattatgcag 4097
Db 5395 GTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATGCAG 5336

QY 4098 tgctgccataaacctagagtataaacactgcggccaacttacttctgacaaacgatcgagg 4157
Db 5335 TGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAACGATCGGAGG 5276

QY 4158 accgaaggagctaaccgcgttttttgcacaacaatg9ggggatcatgtaactcgccttgatcg 4217
Db 5275 ACCGAAGGAGCTAACCGCTTTTGTGCAACAATGGGGGATCATGTAACTCGCCTTGATCG 5216

QY 4218 ttgggaaccggagctgaatgaagccataccaacacgacgagcgtgacaccacgatgcctgt 4277
Db 5215 TTGGGAACCGGAGCTGAATGAAGCCATATCAACACGACGAGCGTGACACCACGATGCCTGT 5156

QY 4278 agcaatggcaacaacgttgcgcaaaactattaactggcgaaactactactactgttcccg 4337
Db 5155 AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCG 5096

QY 4338 gcaacaattaatagactggatggaggcggataaaagtgcaggaccacttctgcgctcggc 4397
Db 5095 GCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACCTTCTCGCTCGGC 5036

QY 4398 ccttcggcgtgctggtttattgctgataaaatctggagccggtgagcgtgggtctcgcg 4457
Db 5035 CCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGTCTCGCGG 4976

QY 4458 tatcattgcagcactggggccagatggtgaagccctcccgtatcgtagttatctacacgac 4517
Db 4975 TATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC 4916

QY 4518 ggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagtgccctcaact 4577
Db 4915 GGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT 4856

QY 4578 gattaagcattggttaactgtcagaccaaagttaactcatatatactttagattgatttaa 4637
Db 4855 GATTAAGCATTGGTAACCTGTGAGACCAAGTTTACTCATATATACTTTTAGATTGATTTAAA 4796

QY 4638 acttcatttttaattaaaaggatctaggtgaagatcccttttgataaatctcatgaccaa 4697
Db 4795 ACTTCATTTTAAATTAAGAGATCTAGGTGAAGATCCTTTTGTGATAATCTCATGACCAA 4736

QY 4698 aatcccttaacgtgagtttctgctccactgagcgtcagacccccgtagaaaaagatcaaaag 4757
Db 4735 AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGGTGACACCCCGTAGAAAAAGATCAAAGG 4676

QY 4758 atcttcttgagatccctttttctgcgcgtaactctgctgcttgcaaacacaaaaaacacc 4817
Db 4675 ATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACCCACC 4616

QY 4818 gctaccagcgggtgtgttgcggatcaagagctaccaactcttttcgaaggtaac 4877
Db 4615 GCTACCAGCGGTGTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCGAAGGTTAAC 4556

QY 4878 tggcttcagcagagcagatatacacaatactgtccttctagtgtagcgcgtagttaggcca 4937
Db 4555 TGGCTTCAGCAGAGCGCAGATACCAAAATACTGCTCTCTAGTGTAGCCGTAGTTAGGCCA 4496

QY 4938 ccacttaagaactctgtagcaccgcctacatacatcgtcgtctgtaactcctgttaccagt 4997
Db 4495 CCACCTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCTGCTAATCTCTTACCAGT 4436

QY 4998 ggctgctgccagtgccgataagtcgtgtcttaccggggttggaactcaagacgatagttacc 5057
Db 4435 GGCTGCTGCCAGTGGCGATAAGTCGTGCTTACC GGTTGGACTCAAGACGATAGTTACC 4376

QY 5058 ggataaggcgagcggtcggggctgaacggggggttcgtgcacacagcccagcttggagcg 5117
Db 4375 GGATAAGGCGCAGCGTTCGGGCTGAACGGGGGTTTCGTGCACACAGCCCAGCTTGGAGCG 4316

QY 5118 aacgacctacacccgaactgagatatacctacagcgtgagctatgagaaaagcgcacgcttcc 5177
Db 4315 AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCCACGCTTCC 4256

QY 5178 cgaaggagaaaaagcgagacaggtatcccgtaagcggcagggctcgaaacagagagcgcac 5237
Db 4255 CGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGGGTTCGAAACAGGAGCGCAC 4196

QY 5238 gagggagcttccaggggaaaaacgcctggtatctttagtccctgctcgggttccgccaact 5297
Db 4195 GAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCTCTGTCCGGTTTCGCCACCT 4136

QY 5298 ctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatgaaaaaacgc 5357
Db 4135 CTGACTTGAGCGTCGATTTTGTGATGCTCGTCAAGSGGGCGGAGCCTATGGAANAACGC 4076

QY 5358 cagcaacgcggccttttttacgggttcctggccttttgccttttgcctcacaatgttctt 5417
Db 4075 CAGCAACGCGGCTTTTACGGTTCCTGGCCTTTTGGTGGCCTTTTGTCTCACATGTTCTT 4016

QY 5418 tcctgcgttatccctgattctgtgataaacctgattaccgccttttgagtgagctgatac 5477
Db 4015 TCCTGCGTTATCCCTGATCTGTGGATAACCCGTATTACCGCCTTTGAGTGAGTGATAC 3956

QY 5478 cgctgcgcgagccgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcag 5537
Db 3955 CGCTCGCCGACGCGAAGACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGAGAGCG 3896

QY 5538 cccaatacgcaaacccctctccccgcggttgccgattcattaatgcagctggcacga 5597
Db 3895 CCCAATACGCAAAACCGCTCTCCCGCGCTTGCCCGATTTCATTAATGCAGCTGCGCACGA 3836

QY 5598 caggtttcccgactgaaaagcgggagcagcagcagcagcagcagcagcagcagcagcagc 5657
Db 3835 CAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAC 3776

QY 5658 tcattaggcacccccaggcctttacacttttatgcttccggcctcgtatgtgtggaattgt 5717
Db 3775 TCATTAGGCACCCCGAGCTTTACACTTTATGCTTCGGCTCGTATGTTGTGTGGAATTGT 3716

QY 5718 gagcgataaacaatttcacacaggaacagctatgaccatgattacgcca 5767
Db 3715 GAGCGGATAACAAATTTACACAGGAACAGCTATGACCATGATTACGCCA 3666

RESULT 3

US-08-789-333F-58/c
; Sequence 58, Application US/08789333F
; Patent No. 6153380
; GENERAL INFORMATION:
; APPLICANT: No. 6153380an, Garry P
; APPLICANT: Rothenberg, S. M.
; TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
; TITLE OF INVENTION: EFFECTOR PEPTIDES AND RNA MOLECULES
; FILE REFERENCE: A642601DJBRMSDSS
; CURRENT APPLICATION NUMBER: US/08/789,333F
; CURRENT FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: 08/589,108
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: 08/589,911
; PRIOR FILING DATE: 1996-01-23
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 4950
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: retroviral

; OTHER INFORMATION: vector with presentation construct sequence.
US-08-789-333F-58

Query Match		45.6%	Score 2628;	DB 3;	Length 4950;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2628;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	3137	cactggcgcgtcgcttttacaaacgctgctgactgggaaaaacccctggcgcttaccccaacttaatc	3196		
Db	4675	CACtGGCGGTCGtTTTACAAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACtTTAAATC	4616		
Qy	3197	gccttgacgacacatcccccttttcgacagctggcgtaataagcgaagagcccgacccgac	3256		
Db	4615	GCCTTGcAGCACATCCCCCTTTCCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCCGATC	4556		
Qy	3257	gcccttcccaacagttgcgcagcctgaatggcgcaatggcgccctgactgcggtattttctcc	3316		
Db	4555	GCCCTTCCCAACAGTTGCGCAGCCCTGAATGGCGAATGGCGCCTGATCGGGTATTtTtCTCC	4496		
Qy	3317	ttacgcatctgtgcggtattttcacaccgcatatggtgcactctcagtcacaatctgctctg	3376		
Db	4495	TTACGCGATCTGTGCGGTATtTCACACCGCATATGGTGCACtCTCAGTACAATCTGCTCTG	4436		
Qy	3377	atgccgcatagttaagccagcccgacacaccccgccaacaccccgctgacgcgcctgacggg	3436		
Db	4435	ATGCCGcATAGTTAAGCCAGCCCCCGACACCCGCCAACACCCCGCTGACGGCCCTGACGGG	4376		
Qy	3437	cttgtctgctcccggcacatccgcttacagacaagctgtgacccgtctccgggagctgcatgt	3496		
Db	4375	CtTTGTCTGCTCCCGGCATCCGGCTTACAGACAAGCTGTGACCCGTCTCCGGGAGCTGCATGT	4316		
Qy	3497	gtcagaaggttttcaccgtcatcacgaaacgcgcgagacgaagggccctcgtgatacgcc	3556		
Db	4315	GTCAGAGGTTTTCACCGTCATCACGAAACGCGGAGACGAAAGGGCCCTCGTGATACGCC	4256		
Qy	3557	tattttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttc	3616		
Db	4255	TATTTTtATAGGTtAATGTcATGATAAATAGTtTCTTAGACGTcAGGTGGCActTTtTC	4196		
Qy	3617	ggggaaatgtgcgcggaacccctattgtttatttttctaaatcatcattcaaatatgtatc	3676		
Db	4195	GGGAAATGTGCGCGGAACCCCTATtTGTTATtTTCTAAATACATtCAAATATGTATC	4136		
Qy	3677	cgctcatgagacaataaacctgataaaatgcttcaataatatgaaaaaggaagagtatga	3736		
Db	4135	CGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATtGAAAAAGGAAGATATGA	4076		
Qy	3737	gtattcaacatttccgctgtgcgcccttatcccttttttgcggcaattttgccttccctgtt	3796		
Db	4075	GTATtCAACATtTCCGTGTGCGCCCTATtTCCCTTTtTTGCGGCATtTTGcCTTCCCTGTTT	4016		
Qy	3797	ttgctcaccagaaacgctggtgaagtaaaaagatgctgaaagatcagttgggtgcacgag	3856		
Db	4015	TTGCTCACCCAGAAACGCTGGTGAAGTAANAAGATGCTGAAGATCAGTTGGGTGCACGAG	3956		
Qy	3857	tgggttacatcgaaactggatctcaacagcgggtaagatcccttgagagtttgcgccccgaag	3916		
Db	3955	TGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCCTTGAGAGTtTTGCCCCGAAG	3896		
Qy	3917	aacgtttttccaatgatgagcacacttttaagttctgctatgtggcgcggtattatcccgt	3976		
Db	3895	AACGTtTTCCAATGATGAGCACtTTTAAAGTtCTGCTATGTGGCGCGGTATTATCCCGTA	3836		
Qy	3977	ttgacgcggggcaagagcaactcggctgcgcgcatacactattctcagaatgacttgggtg	4036		
Db	3835	TTGACGCGCGGCAAGAGCAACTCGGTCGCGGCATACACTATTCTCAGAATGACTTGGTTG	3776		
Qy	4037	agtactcacaggtcacagaaaaagcatcttacggatggcatgacagtgaagagaattatgca	4096		
Db	3775	AGTACTCACAGTcACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATtATGCA	3716		
Qy	4097	gtgctgccataaaccatgagtataaacactgcggccaacttacttctgacaacgacgcggag	4156		

Db	3715	GTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAG	3656		
Qy	4157	gaccgaaggagctaaaccgcttttttgacaacaatggtgggatactgtaaactgccttgatc	4216		
Db	3655	GACCGAAGGAGCTAAACCGCTTTTtTGCAACATGGGGATCATGTAACTGCCTTGATC	3596		
Qy	4217	gttggaaaccggagctgaatgaagccataccaaaacgacgagcgtgacaccacgatgcctg	4276		
Db	3595	GTTGGAAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGATGCCTG	3536		
Qy	4277	tagcaatggcaacaacgttgcgcaaaactattaaactggcgaactacttactctagcttccc	4336		
Db	3535	TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC	3476		
Qy	4337	ggcaacaattaatagactggatggagggcggtataaagtgcaggaccacttctgcctcgg	4396		
Db	3475	GGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACtTCTCGCTCGG	3416		
Qy	4397	cccttcgggctggctggtttattgctgataaaatctggagccggtgagcgtgggtctcgcg	4456		
Db	3415	CCCTTCGGCTGGCTGGTtTATTGTGTATAAAATCTGGAGCCGCTGAGCGTGGGTCTCGCG	3356		
Qy	4457	gtatcattgcagcactggggccagatggttaagccctcccgtatcgtagttatctacacga	4516		
Db	3355	GTATCATTTGCAGCACTGGGGCCAGATGCTAAGCCCTCCCGTATCTAGTtTATCTACACGA	3296		
Qy	4517	cggggagtccaggcaactatggatgaacgaataagacagatcgctgagataggtgcctcac	4576		
Db	3295	CGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGTGCCTCAC	3236		
Qy	4577	tgattaaagcattggtaactgtcagaccgaagtctactcatatatacttttagattgatttaa	4636		
Db	3235	TGATTAAGCATtGGTAACtGTcAGACCcAAAGTtTACTCATATATACTTTAGATtGATTAA	3176		
Qy	4637	aacttcatttttaatttaaaggatctaggtgaagatccctttttgataaatctcatgacca	4696		
Db	3175	AACtTCAATtTTAAATtTAAAGGATCTAGGTGAAGATCCtTTTtTGATAATCTCATGACCA	3116		
Qy	4697	aaatcccttaacgtgagtttctgtccactgagcgtcagacccccgtagaanaagatcaaa	4756		
Db	3115	AAATCCCTTAACGTGAGTtTTCGTTCCACTGAGCGTcAGACCCCGTAGAAAAGATCAAA	3056		
Qy	4757	gatcttctgagatccctttttctgcggtaactgtcgtgcttgcaacaaaaaacccac	4816		
Db	3055	GATCTTCTTGAGATCCtTTTtTCTGCGGTAAATCTGCTGCTTGCAACAACAAAAACCAC	2996		
Qy	4817	cgctaccagcgggtggttgttgcgggatcaagagctaccaaactctttttccgaaggtaa	4876		
Db	2995	CGTACCAGCGGTGGTtTGTtTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA	2936		
Qy	4877	ctggcttcagcagagcgcagatataccaaaatactgtccttcttagtgcgctagttagcc	4936		
Db	2935	CTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTTAGTGTAGCCGTAGTTAGGCC	2876		
Qy	4937	accacttcaagaactctgtagcaccgcctacataacctcgtctctgtaaatcctgttaccag	4996		
Db	2875	ACCACtTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAACTCTGTtTACCAG	2816		
Qy	4997	tggctgtccagtgggcgataaagtctgtcttaccgggttggaactcaagacgatatgac	5056		
Db	2815	TGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTtTGACTCAAGACGATAGTTAC	2756		
Qy	5057	cggataaggcgcgagcggctcgggctgaacgggggggttctgtgcacacagcccagcttggagc	5116		
Db	2755	CGGATAAGGCGCAGCGGTCCGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGAGC	2696		
Qy	5117	gaacgacctacacccgaactgagataccctacagcgtgagctatgagaaagcgccacgcttc	5176		
Db	2695	GAACGACCTACACCGAACTGAGATACCTACACGCTGAGCTATGAGAAAAGCCACCGCTTC	2636		
Qy	5177	ccgaaggggagaaaaggcggaacaggatatccggtaagcggcagggctcggaacagagagcga	5236		

Db 8737 AGTACTACCACTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCA 8678
QY 4097 gtgctgccataaccatgagtataaactgcggccaaacttacttctgacaacgacgcgag 4156
Db 8677 GTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAG 8618
QY 4157 gaccgaaggagctaaccgccttttttgcacaacatggggatcatgtaactcgccttgatc 4216
Db 8617 GACCGAAGGAGCTAACCGCTTTTGTGACACATGCGGGATCATGTAACTCGCCTTGATC 8558
QY 4217 gttgggaaccgagctgaatgaagcccatatacacaacgacgagcgtgacaccaagatgcctg 4276
Db 8557 GTTGGGAACCCGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACACGATGCCTG 8498
QY 4277 tagcaatggcaacaacgcttgcgcaaaactataaactggcgaaactacttactctagcttccc 4336
Db 8497 TAGCAATGGCAACAACGTTGCGCAAACTATAACTGGCGAACTACTTACTCTAGCTTCCC 8438
QY 4337 ggcaacaattaatagactggatggagggcggaataaaagttgcaggaccacttctgcgctcgg 4396
Db 8437 GGCAACAATTAAATAGACTGGATGGAGGGGATAAAAGTTGCAGGACCACACTTCTGCGCTCGG 8378
QY 4397 cccttcggctggctggtttattgctgataaaatctggagccggtgagcgtgggtctcgcg 4456
Db 8377 CCCTTCCGGCTGGCTGTTTATTGCTGATAAATCTGGAGCCGTTGAGCGTGGGTCTCGCG 8318
QY 4457 gtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacga 4516
Db 8317 GTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCCTATCGTAGTTATCTACACGA 8258
QY 4517 cggggagtcaggcaactatggatgaacgaataagacagatcgcgtgagataggtgcctcac 4576
Db 8257 CGGGGAGTCAGGCAACTATGGATGAACGAATAAGACAGATCGCTGAGATAGGTGCCTCAC 8198
QY 4577 tgattaagcatggttaactgtcagaccagtttactcatatatactttagattgattaa 4636
Db 8197 TGATTAAGCATTTGTAACCTGTGAGACCAAGTTTACTCATATATACTTTAGATTGATTAA 8138
QY 4637 aacttcatttttaatttaaaggatctaggtgaagatcccttttggataaatctcatgacca 4696
Db 8137 AACTTTCATTTTAAATTTAAAGGATCTAGTGGAAGATCCCTTTTGTGATAATCTCATGACCA 8078
QY 4697 aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaagatcaaa 4756
Db 8077 AAATCCCTTAAAGTGAGTTTTCGTTCCACTGAGCGTGCAGACCCCGTAGAAAAAGATCAAAG 8018
QY 4757 gatcttcttgagatccctttttctgcggtaaatctgctgcttgcaacaaaaaaaccac 4816
Db 8017 GATCTTCTTGAGATCCTTTTTCGTCGCGGTAATCTGCTGCTTGCAAAACAAAACCAC 7958
QY 4817 cgctaccagcgggtggtttgttgcggatcaagagactaccaactcttttccgaaggtaa 4876
Db 7957 CGCTACCAGCGGTGTTGTTTGGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAA 7898
QY 4877 ctggcttcagcagagcgcagataccataactgtccttctagtgtagcgtagttaggcc 4936
Db 7897 CTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTTAGTGTAGCGTAGTTAGGCC 7838
QY 4937 accacttcagaactctgtagcacccgctacatacctcgtctgtctaactcgtttaccag 4996
Db 7837 ACCACTTCAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAG 7778
QY 4997 tggctgctgccagtggcgataaagtctgtcttaccggggttggactcaagacgatatgtac 5056
Db 7777 TGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACC GGTTGGACTCAAGACGATAGTTAC 7718
QY 5057 cggataaaggcgcagcggctcgggctgaacgggggggttcgtgcacacagcccagcttgagc 5116
Db 7717 CGGATAAGCGCAGCGGTGCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGC 7658
QY 5117 gaacgacctacaccgaactgagataacctacagcgtgagctatgagaaagcgccacgcttc 5176
Db 7657 GAACGACCTACACCGAACTGAGATACCTACAGCGTGAAGCTATGAGAAAGCGCCACGCTTC 7598

QY 5177 ccgaaggagaaaggcgacaggtatccggttaagcggcagggtcggaaacagagagcgca 5236
Db 7597 CCGAAGGGAGAAAGCGCGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGCGCA 7538
QY 5237 cgaggagcttccaggggggaaacgcctggtatctttatagtcctgtcgggttcgccacc 5296
Db 7537 CGAGGGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTCTGTCGGGTTTCGCCACC 7478
QY 5297 tctgacttgagcgtcgatttttctgagctcgtcagggggcgagccctatgaaaaaacg 5356
Db 7477 TCTGACTTGAGCGTCGATTTTGTGATGCTCTCGTCAGGGGGCGGAGCCCTATGGAACACG 7418
QY 5357 ccagcaacgcggcccttttttacgggttcctggcccttttctggcccttttgcacatgttct 5416
Db 7417 CCAGCAACGCGGCCTTTTACGGTTCCTGGCCCTTTTGCTGCATGTTCT 7358
QY 5417 ttctcgcttatccctgattctctgtggataaacgctattaccgcctttgagtgagctgata 5476
Db 7357 TTCTCTGGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTTGAGTGAGCTGATA 7298
QY 5477 ccgctcgcgcagccgaacgacgacgagcgcagcagcgttgggccgattcattaaatgcagctggcacg 5536
Db 7297 CCGCTCGCGCAGCGCAACGACCGAGCGCAGCTCAGTGAGCGAGGAAGGAGC 7238
QY 5537 gcccaatacgcgaacgcctctcccgcgcgttggccgattcattaaatgcagctggcacg 5596
Db 7237 GCCCAATACGCAACCGCCTCTCCCCGCGCGTTGCCGATTCATTAAATGCAGTGGCACC 7178
QY 5597 acaggtttcccgactggaaagcgggcagtgagcgcaacgcaataatgtgagttagctca 5656
Db 7177 ACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTAGCTCA 7118
QY 5657 ctcattaggcacccaggttttacactttatgcttcggcgtcgtatgttgtggaattg 5716
Db 7117 CTCATTAGGCACCCAGGCTTTACACTTTATGCTCCGGCTCGTATGTTGTGTGAATTG 7058
QY 5717 tgagcgataaacaatttcacacaggaacagctatgaccatgattacg 5764
Db 7057 TGAGCGGATAACAATTTACACACAGGAAACAGCTATGACCATGATTACG 7010

RESULT 5
US-08-343-401A-3/c
; Sequence 3, Application US/08343401A
; Patent No. 5661132
; GENERAL INFORMATION:
; APPLICANT: Swain, William F
; APPLICANT: Macklin, Michael D
; APPLICANT: Eriksson, Elof
; APPLICANT: Andree, Christophe
; TITLE OF INVENTION: Improved Wound Healing
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: PO Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,401A
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 11-229-9103-9


```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: PWRG1630
; FEATURE:
; NAME/KEY: exon
; LOCATION: 713..721
; FEATURE:
; NAME/KEY: exon
; LOCATION: 981..1253
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1253)
;
; NAME/KEY: sig_peptide
; LOCATION: 713..1049
;
US-08-343-401A-3
```

```
Query Match 43.9%; Score 2533; DB 1; Length 4283;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3133 aattcactggccgtcggttttacaaacgtcgtagctgggaaaaaccctggcggtaccacaactt 3192
|||||
Db 4265 AATTCACTGGCCGTCGTTTACAAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTT 4206

QY 3193 aatcgccctgcagcacatcccccttccgccagctggcgtaatagcggaagggccgcacc 3252
|||||
Db 4205 AATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCGCCACC 4146

QY 3253 gatcgccctcccaacagtgcgcagcctgaatggcgaaatggcgctgatgcggtatttt 3312
|||||
Db 4145 GATCGCCCTTCCCAACAGTTGCGAGCCTGAATGGCGAATGGCGCTGATGCGGTATTT 4086

QY 3313 ctccctacgcactgtcggttatctcacaccgcataatggtgcactctcagtaacaatctgc 3372
|||||
Db 4085 CTCCTTACGCATCTGTGCGGTATTTACACACCGCATATGGTGCACTCTCAGTACAATCTGC 4026

QY 3373 tctgatccgcatagttaagccagccccgcacccgcgaacaccccgctgacgcgcccctga 3432
|||||
Db 4025 TCTGATGCCGCATAGTTAAGCCAGCCCCCGACACCCGCCAACACCCCGTGACCGCTGAC 3966

QY 3433 cgggcttgctgctccggcatccgcttacagacaagctgtgacccgtctccgggagctgc 3492
|||||
Db 3965 CGGGCTTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC 3906

QY 3493 atgtgtcagaggttttccaccgtcatcacccgaacgcgcgagacgaaagggcctcgtgata 3552
|||||
Db 3905 ATGTGTCAGAGGTTTTACACCGTCATCACCGAACCAGCGGAGACGNAAGGGCCTCGTGATA 3846

QY 3553 cgcctattttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcact 3612
|||||
Db 3845 CGCCTATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACT 3786

QY 3613 ttccggggaaatgtgcgcggaaacccctatttgtttatttttctaaatacatctcaaatatg 3672
|||||
Db 3785 TTTCGGGGAATGTGCCGGAAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATG 3726

QY 3673 tatccgctcatgagacaaataacccttgataaaatgcttcaataataattgaaaaaggaagtg 3732
|||||
Db 3725 TATCCGCTCATGAGACAAATAACCCTGATAAATGCTCAATAATATTGAAAAGGAAGAGT 3666

QY 3733 atgagtattcaacatttcgcgtgtcgcccttattcccttttttgcggcattttgccttcct 3792
|||||
```

```
Db 3665 ATGAGTATTCAACATTTCCGTGTCCGCCCTTATTCCCTTTTTTTCGGGCATTTTGCCCTTCCT 3606

QY 3793 gtttttgcaccagaaacgcctggtgaaagtataaagatgctgaagatcagttgggtgca 3852
|||||
Db 3605 GTTTTGTCTACCCACCAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 3546

QY 3853 cgagtgggttacatcgaaactggatctctcaacagcggtaagatcccttgagagttttcgcccc 3912
|||||
Db 3545 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC 3486

QY 3913 gaagaacgttttccaatgatgagcaccttttaaagttctgctatgtggcggtattatcc 3972
|||||
Db 3485 GAAGAACGTTTTTCCAATGATGAGCACTTTTTAAAGTTCTGCTATGTGGCGGTTATTATCC 3426

QY 3973 cgtattgacgcgggcaagagacaactcggtcgcgcgcatacacactattctcagaatgacttg 4032
|||||
Db 3425 CGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTG 3366

QY 4033 gttgagtactcaccagtcacagaaaaagcatcttaccggatggcatgacagtaagagaatta 4092
|||||
Db 3365 GTTGAGTACTCACCACTCACAGAAAAGCATCTTTACGGATGGCATGACAGTAAGAGAATTA 3306

QY 4093 tgcagtgtgccataaccatgagtgtataaacactcggcgccaaacttacttctgcacaacgac 4152
|||||
Db 3305 TGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATC 3246

QY 4153 ggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccctt 4212
|||||
Db 3245 GGAGGACCGAAGGAGCTAAACCGCTTTTTTGTGCACAACATGGGGATCATGTAACTCGCCTT 3186

QY 4213 gatcgttgggaaccggagctgaatgaagccataccaaaaagcagcgtgacaccacgatg 4272
|||||
Db 3185 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACACGATG 3126

QY 4273 cctgtagcaatggcaacaacgttgcgcaaaactattaaactggcgaacttacttactctagct 4332
|||||
Db 3125 CCTGTAGCAATGCCRAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCT 3066

QY 4333 tcccggcaacaattaatagactggatggagggcggaataaaagtgcaggaccacttctgcgc 4392
|||||
Db 3065 TCCCGGCAACAATTAATAGACTGGATGGAGGGCGGATAAAGTTGCAGGACCACCTCTGCGC 3006

QY 4393 tcggcccttcggctggctggtttattgctgtataaaatctggagccggtgagcgtgggtct 4452
|||||
Db 3005 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGTGATAAATCTGGAGCCGGTGAGCGTGGGTCT 2946

QY 4453 cgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagtattctac 4512
|||||
Db 2945 CGCGGTATCATTCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC 2886

QY 4513 acgacggggagtcaggcaactatggatgaacgaataagacagatcgctgagataggtgcc 4572
|||||
Db 2885 ACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC 2826

QY 4573 tcaactgattaaagcattgggtaactgtcagaccaagtttactcatatatactttagattgat 4632
|||||
Db 2825 TCACTGATTAAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT 2766

QY 4633 ttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttbgataaatctcatg 4692
|||||
Db 2765 TTAAAACTTCATTTTAAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATG 2706

QY 4693 accaaaaatcccttaacgtgagtttcttcgtccactgagcgtcagacccccgtagaaaaagatc 4752
|||||
Db 2705 ACCAAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATC 2646

QY 4753 aaaggatctcttgagatccctttttctgcgcgctaatctgctgcttgcaacaaaaaaa 4812
|||||
Db 2645 AAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAANA 2586

QY 4813 ccaccgctaccagcgggtggtttgttgcgggatcgaagactaccacactcttttccgaag 4872
|||||
Db 2585 CCACCGCTACCAGCGGTGTTGTTTGGCCGGATCAAGAGCTACCAACTCTTTTTCGGAAG 2526
```


QY 4873 gtaactggcttcagcagagcgagataccaaatactgtccttcttagttagccgtagtta 4932
|||||
Db 2525 GTAACGTGGCTTCAGCAGAGCGCAGATACCAAATACTGCTCTTAGTGTAGCCGTAGTTA 2466

QY 4933 ggcacaccacttcaagaactctgtagcaccgcctacatacctcgtctctgtaaatcctgta 4992
|||||
Db 2465 GGCACCACTTCAAGAACTCTGTAGCACCCGCCCTACATACCTCGCTCTGCTAATCCTGTTA 2406

QY 4993 ceagtggctgctgccagtgggcgataaagtcgtgtcttaccgggttgactcaagacgatatag 5052
|||||
Db 2405 CCAGTGGCTGCTGCCAGTGGCGATAAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAG 2346

QY 5053 ttaccggataaggcgcgagcgtcgggctgaacgggggttcgtgcacacagcccagcttg 5112
|||||
Db 2345 TTACCGGATAAGGCGCAGCGCTCGGGCTGAACGGGGGTTTCGTGCACACAGCCCAGCTTG 2286

QY 5113 gaggaacgacctacaccgaactgagatatacctacagcgtgagctatgagaaaagcgccacg 5172
|||||
Db 2285 GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACG 2226

QY 5173 ctccccgaaggagaaaggcgacaggtatccggtaagcggcaggggtcggaaacaggagag 5232
|||||
Db 2225 CTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAG 2166

QY 5233 cgcagcagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggttcgc 5292
|||||
Db 2165 CGCACGAGGGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTGTTCGGGTTTCG 2106

QY 5293 cacctgacttgagcgtcgatttttgtgatcgtcgagggggcgagcctatggaaa 5352
|||||
Db 2105 CACCTGTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATCGNAA 2046

QY 5353 aacgcagcaacgcggcctttttacgggttcctggccttttgcgtggccttttgcgcacatg 5412
|||||
Db 2045 AACGCAGCAACGCGGCCCTTTTACGGTTCCTGGCCCTTTTGTGGCCCTTTTGTCTCACATG 1986

QY 5413 ttctttcctgcgttatccccctgattctgtggataaacctgattaccgcctttgagtgagct 5472
|||||
Db 1985 TTTCTTTCTCGCTTATCCCCCTGATCTGTGGATAACCCGTATACCGCCCTTTGAGTGAGCT 1926

QY 5473 gatacgcgtcgcgcagcgcgaacgacgcagcgcagcgcagtcagtcagtgagcaggaagcgga 5532
|||||
Db 1925 GATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAA 1866

QY 5533 gagegcccaatacgaaaaccgcctctccccgcgcgttggcgcgatttcattaatgcagctgg 5592
|||||
Db 1865 GAGCGCCCAATACGCNAACCGCCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGG 1806

QY 5593 cacgacaggtttcccgactggaaagcggcgagtcagtcagtcagtcagtcagtcagtcagtcag 5652
|||||
Db 1805 CACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAG 1746

QY 5653 ctactcattaggcacccccaggctttacactttatgcttcgggctcgtatgttggtgga 5712
|||||
Db 1745 CTCACATATTAGGCACCCAGGCTTACACTTTATGCTTCCGCTCGTATGTTGTGTGA 1686

QY 5713 attgtgagcggataaacaatttcacacaggaaacagcgtatgaccatgattacgcca 5767
|||||
Db 1685 ATTTGAGCGGATAACAAATTTACACAGGAACAGCTATGACCATGATTACGCCA 1631

RESULT 6

US-08-445-265A-1/c
; Sequence 1, Application US/08445265A
; Patent No. 5697901
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Elof
; TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street

; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,265A
; FILING DATE:
; CLASSIFICATION: 604
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 110229.91080
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Plasmid DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1250)
US-08-445-265A-1

Query Match 43.9%; Score 2533; DB 1; Length 4283;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3133 aattcaactggccgtcgtttttacaacgtcgtgactgggaaaaacccctggcgttacccaaactt 3192
|||||
Db 4265 AATTCACTGGCGCTCGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTT 4206

QY 3193 aatgccttgccagcacatcccccttttcgccagctggcgttaataagcgaagagccccgcacc 3252
|||||
Db 4205 AATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCCGCACC 4146

QY 3253 gatgccttccccaaacagttggcgcagcctgaatggcgaatggcgctgatcggtattttt 3312
|||||
Db 4145 GATCGCCCTTCCCAACAGTTGGCGACGCTGAATGGCGAATGGCGCTGATGCGGTATTTT 4086

QY 3313 ctccctaacgcattcgtgcggtattttcaacccgcataatggtgcactctcagtcacaattctgc 3372
|||||
Db 4085 CTCCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTTCAGTACAATCTGC 4026

QY 3373 tctgatgcgcgcatagttaagccagcccccgcagacacccgcaacacccgcctgacgcgccctga 3432
|||||
Db 4025 TCTGATGCCGCATAGTTAAGCCAGCCCGCGACACCCGCCAACACCCGCTGACGCGCCCTGA 3966

QY 3433 cgggcttgctgctccccggcattccgcttacagacagaagctgtgacggtctccgggagctgc 3492
|||||
Db 3965 CGGGCTTGCTGCTCCCCGCATCCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC 3906

QY 3493 atgtgtcagaggtttttcacccgtcatcacccgaaacgcgcgcagagacgaaaaggccctcgtgata 3552
|||||
Db 3905 ATGTGTCAGAGGTTTTCACCGTCTATACCGAAACCGCGGAGACGAAAGGGCCCTCGTGATA 3846

QY 3553 cgcctatttttataggttaagtcatgataataatggtttcttagcagtcagggtgcact 3612
|||||
Db 3845 CGCCTATTTTATAGGTTAATGTCAATATAATGTTTCTTAGAGCTCAGGTGGCACT 3786

QY 3613 ttccggggaaaatgtgcgcggaacccctatttgtttatttttctaatacatcattcaaatatg 3672
|||||
Db 3785 TTTCGGGGAAATGTGCGCGGAACCCCTATTGTTTATTATTCTTAATAACATTCAAATATG 3726

Qy	3673	tatccgctcatgagacaataaacccctgataaaatgcttcaataataattgaaaaaggaagagt	3732
Db	3725	TATCCGCTCATGAGACANATNACCCTGATAAATGCTTCAATAATATGAAAAAGGAAGAGT	3666
Qy	3733	atgagtattcaacattccggtgcgccttattcccttttttgcggcattttgccttcct	3792
Db	3665	ATGAGTATTCAACATTTCCGTTGTCGCCCTTATTCCTTTTTCGCCGATTTTGCTTCCCT	3606
Qy	3793	gttttttgcaccaccagaaaacgctggtgaaaagtataaagatgctgaagatcagttgggtgca	3852
Db	3605	GT'TTTTGC'TCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGAT'CAGTTGGGTGCA	3546
Qy	3853	cgagtgggttacatcgaaactggatctcaacagcggtaagatccttgagagtttttcgcccc	3912
Db	3545	CGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCC	3486
Qy	3913	gaagaacgttttccaatgatgagcacttttaaaagtctgctatgtggcgcggtattatcc	3972
Db	3485	GAAGAACGTTTTCGAATGATGAGCACTTTTAAAGTTCTGCTATGGCGCGGTATTATCC	3426
Qy	3973	cgattgacccggggaagagcaactcggtcgccgcatacactattctcagaatgacttg	4032
Db	3425	CGTATTGACCGCGGGAAGAGCAACTCGTGGCGGCATACACTATTCTCAGAATGACTTG	3366
Qy	4033	gttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagaaatta	4092
Db	3365	GT'TGAGTACTCACCAAGTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGATTA	3306
Qy	4093	tgcagtgctgccataaaccatgagtataaacactgcggccaacttacttctgcacaacgatc	4152
Db	3305	TGCAGTGTGCCATAAACCATGAGTGATAAACACTGCGGCCAACTTACTTCTGACAAACGATC	3246
Qy	4153	ggaggaccggaaggagctaaccgccttttttgcacaacatgggggatcatgtaaactgcctt	4212
Db	3245	GGAGGACCGAAGGAGCTAAACCGCTTTTTCACAAACATGGGGGATCATGTAAC'TCGCCTT	3186
Qy	4213	gatcgttgggaaccggagctgaatgaagccatacaciaaacgacgagcgtgacacacgagt	4272
Db	3185	GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACGATG	3126
Qy	4273	cctgtagcaatggcaacaacgttgcgaaaactattaactggcgaaactacttactctagct	4332
Db	3125	CCTGTAGCAATGCCAAACAACGTTGCGCAAAC'TATTAACTGGCGAACTACTTACTCTAGCT	3066
Qy	4333	tcccggaacaattaatagactggatggaggcggaataaagttgcaggaccacttctgcgc	4392
Db	3065	TCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGACCACACTTCTGCGC	3006
Qy	4393	tcggcccttcggcgctggctgtttattgctgataaaactggagccgggtgagcgtgggtct	4452
Db	3005	TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCT	2946
Qy	4453	cgcggtatcaattgcagcactggggccagatggttaagccctcccgtatcgtagtattctac	4512
Db	2945	CGCGGTATCATTCAGCACACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC	2886
Qy	4513	acgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcc	4572
Db	2885	ACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC	2826
Qy	4573	tcactgattaaagcatgtgtaactgtcagaccaaagtttactcatatatacttttagattgat	4632
Db	2825	TCACTGATTAAAGCATTTGGTAAC'TGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT	2766
Qy	4633	ttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttttgataatctcatg	4692
Db	2765	TTAAAAACTTCATTTTAA'TTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATG	2706
Qy	4693	accaaaaatcccttaacgtgagtttttcggttccactgagcgtcagacccccgtagaaaaagatc	4752
Db	2705	ACCAAAATCCCTTAAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGGTAGAAAAGATC	2646

Qy	4753	aaaggatcttcttgagatccctttttttctgcgcgtaaatctgctgcttctgcaaaaaaa	4812
Db	2645	AAAGGATCTTCTTGAGATCCTTTTTCGCGGTAATCTGCTGCTGCAACAAAAA	2586
Qy	4813	ccaccgctaccagcgggtggtttgtttgcgggatcaagagctaccaactctttttccgaag	4872
Db	2585	CCACCGCTACCAGCGGTGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGAAG	2526
Qy	4873	gtaactggcttcagcagagcgcagataccaaaaactgtccttctagtgtagccgtagtta	4932
Db	2525	GTAAC'TGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCCGTAGTTA	2466
Qy	4933	ggccaccacttcaagaactctgtagcacgcctacatacctcgctctgctaactcctgta	4992
Db	2465	GGCCACCAC'TTCAAGAACTCTGTAGCACCGCTACATACTCGCTCTGCTAATCTGTTA	2406
Qy	4993	ccagtggctgctgccagtggcgataaagtcgtcttaccgggttggactcaagacgatatg	5052
Db	2405	CCAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACC'GGT'TGGACTCAAGACGATAG	2346
Qy	5053	ttaccggataaaggcgagcggctcgggtgaaacgggggttcgtgcacacagccccagcttg	5112
Db	2345	TTACCGGATAAGGCGCAGCGGTGCGGCTGAAACGGGGGTTCGTGTCACACAGCCCCAGCTTG	2286
Qy	5113	gagcgaacgaactacacccaactgagatacctacagcgtgagctatgagaaagcgccaag	5172
Db	2285	GAGCGAACGACTACACCCGAAC'TGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACG	2226
Qy	5173	cttccccgaaggagaaagcgcgacaggtatccggtaagcgagggtcgggaacaggagag	5232
Db	2225	CTTCCCGAAGGAGAAAAAGCGGACAGGTATCCGGTAAAGCGGAGGTGCGGAACAGGAGAG	2166
Qy	5233	cgcacgaggaggttccaggggggaaaacgcctgggtatctttatagtcctgtcgggttcgc	5292
Db	2165	CGCACGAGGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCTCTCGGGTTTCGC	2106
Qy	5293	cacctctgacttgagcgtcgatcttttctgtagtcgtcagggggcgagcgcttatggaaa	5352
Db	2105	CACCTCTGACTTGAGCGTCGATTTT'TGTGATGCTCGTCAGGGGGCGGAGCCTATCGAAA	2046
Qy	5353	aacgccaagcaacgcggcctttttacgggttccttgcgccttttgcgccttttgcctcacatg	5412
Db	2045	AACGCCAGCAACGGGGCTTTTATACGGTTCTTGGCCTTTTGTCTGGCCTTTTGTCTCACATG	1986
Qy	5413	ttctttcctgctgttatccccctgattctgttggtataaccgtattaccgcctttgagtgagct	5472
Db	1985	TTCTTTCTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACC'GCTTTGAGTGAGCT	1926
Qy	5473	gataccgctcgccgcagccgaacgacgagcgagcgagtcagtcagtcagtcagtcagtcagtcag	5532
Db	1925	GATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTCAGCGAGGAGGAAACGGAA	1866
Qy	5533	gagcgcccaatacgcgaacccgcctctcccgcgcgcttgggccgattcattaatgcagctgg	5592
Db	1865	GAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTGGCCGATTCATTAAATGCAGCTGG	1806
Qy	5593	cacgacaggtttcccgactggaaaagcgggcagtcagcgaacgcaattaatgtgagttag	5652
Db	1805	CACGACAGGTTTCCCGACTGGAAAGCGGCGCAGTGAGCGCAACGCAATTAATGTGAGTTAG	1746
Qy	5653	ctcaactcattaggcaccccaggctttacactttatgcttccggctcgatatgttgtgtgga	5712
Db	1745	CTCACTCATTAGGCACCCCGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGA	1686
Qy	5713	attgtgagcgggataaacaatttcacacaggaacagctatgaccatgattacgcca	5767
Db	1685	ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCTATGACCATGATTACGCCA	1631

RESULT 7
us-08-990-442-1/c
; Sequence 1, Application US/08990442
; Patent No. 6090790

; GENERAL INFORMATION:
; APPLICANT: Eriksson, Elof
; TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,442
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 310558.90028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Plasmid DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1250)
;
US-08-990-442-1

Query Match 43.9%; Score 2533; DB 3; Length 4283;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3133 aattcaactggcgtcggttttacaacgtcgtgactgggaaacccctggcgttaccccaactt 3192
|||||
Db 4265 AATTCACTGGCGTCGTTTTTACAACGTCGTCGACTGGGAAACCCCTGGCGTTACCCAACTT 4206

Qy 3193 aatcgcccttgagcacatcccccttcgccagctggcgtaaatagcgaagagcccgacc 3252
|||||
Db 4205 AATCGCCTTGACGACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCGCACC 4146

Qy 3253 gatcgcccttcccaacagttggcgagccctgaatggcgaaatggcgccctgatgcggtattt 3312
|||||
Db 4145 GATCGCCCTTCCCAACAGTTGGCGAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTT 4086

Qy 3313 ctccctacgcattctgtgcggtatttcaacccgcataatggtgaactctcagtaacaatctgc 3372
|||||
Db 4085 CTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACTCTCAGTACATCTGC 4026

Qy 3373 tctgatccgcatagttaagccagccccgcacacccgcacacccgcctgacgcgccctga 3432
|||||
Db 4025 TCTGATCCGCATAGTTAAGCCAGCCCGGACACCCGCCAACACCCGCTGACGCGCCCTGA 3966

Qy 3433 cgggctgtctgctcccgccatccgcttacagacaagctgtgaccgtctccgggagctgc 3492
|||||
Db 3965 CGGGCTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCTGC 3906

Qy 3493 atgtctcagagggttttcaaccgtcatcacccgaaacgcgcgagagaaagggcctcgtgata 3552
|||||
Db 3905 ATGTGTACAGAGGTTTTTCACCGTCATCACCGAAACGGCGAGACGAAAGGGCCTCGTGATA 3846

Qy 3553 cgcctattttataggttaatgtcatgataataatggtttcttagagctcagggtgcact 3612
|||||
Db 3845 CGCCTATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGCCACT 3786

Qy 3613 ttctcggggaatgtgcgcgggaacccctatttggtttatttttctaaatacattcaaatatg 3672
|||||
Db 3785 TTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTTTCTAATAACATTCAAATATG 3726

Qy 3673 tatccgctcatgagacaataaacccctgataaaatgcttcaataatattgaaaaaggaagat 3732
|||||
Db 3725 TATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAT 3666

Qy 3733 atgagtattcaacatttccggtgcgcccttattcccttttttgcggcattttgccttct 3792
|||||
Db 3665 ATGAGTATTCAACATTTCCGTGTCGCCCTATTCCCTTTTTCGGCATTTTGCCTTCCT 3606

Qy 3793 gttttgtcaccacagaaacgctggtgaaagtataaagatgctgaagatcagttgggtgca 3852
|||||
Db 3605 GTTTTGTCTCACCCAGAAACGCTGGTGAAGATAAAGATGCTGAAGATCAGTTGGGTGCA 3546

Qy 3853 cgagtgggttacatcgaaactggatctcaacagcgggtaagatcccttgagagttttcgcctc 3912
|||||
Db 3545 CGAGTGGGTACATCGAACTGGATCTCACAGCGGTAAGATCCTTTGAGAGTTTTCGCCCC 3486

Qy 3913 gaagaaacgttttccaatgatgagcaacttttaaagttctgctatgtgcgcggtattatcc 3972
|||||
Db 3485 GAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTATTATCC 3426

Qy 3973 cgtattgcgcggggaagagcaactcgtgcgcgcatacacactattctcagaatgacttg 4032
|||||
Db 3425 CGTATTGACGCGGGCAAGAGCAACTCGTGCGCGCATACACATTATCTCAGAATGACTTG 3366

Qy 4033 gttgagtactcaccagtcacagaaaaagcatcttaccggtggtgcatgacagtaagaaatta 4092
|||||
Db 3365 GTTGAGTACTCACCAAGTACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA 3306

Qy 4093 tgcagtgtcgtccataaaccatgagtataaacactgcggccaacttacttctgacaaacgac 4152
|||||
Db 3305 TGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACCTTACTTCTGACAACCGATC 3246

Qy 4153 ggaggaccggaaggagctaaccgcttttttgcacaacatgggggagatcatgtaactcgcctt 4212
|||||
Db 3245 GGAGGACCGAAGGAGCTAACCGCTTTTTCACAACATGGGGGATCATGTAACCTGCCTT 3186

Qy 4213 gatcgttgggaacccggagctgaatgaagcacaataccaaacgacgagcgtgacacacgatg 4272
|||||
Db 3185 GATCGTTGGGAACCCGGAGCTGAATGAAGCATACCAAAACGACGAGCGTGACACCAACGATG 3126

Qy 4273 cctgtagcaatggcaacaacgcttgccgaaactattaaactggcgaactacttactctagct 4332
|||||
Db 3125 CCTGTAGCAATGCCAACCAACGTTTGCGAAACTATTAACTGGCGAACCTACTTACTCTAGCT 3066

Qy 4333 tccccggcaacaattaatagactggatggaggcgggataaaagtgcaggaccacttctgcgc 4392
|||||
Db 3065 TCCCCGGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACTTCTGCGC 3006

Qy 4393 tcggcccttcgggctggctggtttatttctgtgataaaatctggagccggtgagcgtgggtct 4452
|||||
Db 3005 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAACTCTGGAGCCGGTGAGCGTGGGTCT 2946

Qy 4453 cgcggtataattgcagcactggggccagatggtaagccctcccgatatcgtagttatctac 4512
|||||
Db 2945 CGCGGTATCATATGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC 2886

Qy 4513 acgacgggagtcaggcaactatggtgaacgaaatagacagatcgtgagataggtgcc 4572
Db 2885 ACGACGGGAGTCAGGCAACTATGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC 2826

Qy 4573 tcactgattaagcattggtaactgtcagaccaaagtttactcatatatacttttagattgat 4632
Db 2825 TCACGTGATTAAGCATTTGGTAACCTGTCAGACCAAGTTTACTCATATATATTTAGATTGAT 2766

Qy 4633 ttaaaacttcatttttaatttaaaaggatcttaggtgaagatcccttttggataaatctcatg 4692

Db 2765 TTAATACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGATAATCTCATG 2706
QY 4693 accaaaaaccttaacgtgagttttcttcactgagcgtcagaccgcgtagaaaaagatc 4752
Db 2705 ACCAAATCCCTTAACGTGAGTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATC 2646
QY 4753 aaaggatcttcttgagatccctttttctgcgcgtaactcgtcgtctgcaaacaaaaaaa 4812
Db 2645 AAAGGATCTTTTGAGATCCTTTTTCGCGCTAATCTGCTGCTGCAACAAAAA 2586
QY 4813 ccaccgctaccagcgggtggtttgtttgcgggatcaagagctaccaactctttttccgaag 4872
Db 2585 CCACCGCTACAGCGGTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAAG 2526
QY 4873 gtaactggctcagcagcagcagatcaccaataactgtccttctagtagccgtagtta 4932
Db 2525 GTAACGGCTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTA 2466
QY 4933 ggcaccacttcaagaactctgtagcacgcgcctacatacctcgtcgtcgttaactcctgta 4992
Db 2465 GGCCACCACCTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGTAATCCTGTTA 2406
QY 4993 ccagtggtcgtgccagtggcgataagtgtgtcttaccgggttgactcaagacgatag 5052
Db 2405 CCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGTTGGACTCAAGACGATAG 2346
QY 5053 ttaccggataaggcgcagcgggtcgggtaacgggggttcgtgcacacagcccagcttg 5112
Db 2345 TTACCGGATAAGCGCGCAGCGTCCGGCTGAACGGGGGTTTCGTGCACACAGCCCCAGCTTG 2286
QY 5113 gagcgaacacctacacaccgaactgagatacctacacgcgtgagctatgagaaaagcgcacg 5172
Db 2285 GAGGGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACG 2226
QY 5173 ctccccgaaggagaaaggcggacaggtatccggttaagcgcgaggtcggaacaggagag 5232
Db 2225 CTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAAACAGGAGAG 2166
QY 5233 cgcacgaggagcttccagggggaaacgcgtgtatctttatagtcctgtcgtcggtttcgc 5292
Db 2165 CGCACGAGGAGCTTCAGGGGGAACGCCTGGTATCTTTATAGTCTGTCTCGGTTTCGC 2106
QY 5293 cacctctgactgagcgtcgatcttttgatgctcgtcaggggggcgagcctatggaaa 5352
Db 2105 CACCTCTGACTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATCGAAA 2046
QY 5353 aacgccagcaacgcggcctttttacggttctcgtgaccttttgcgtgaccttttgctcacatg 5412
Db 2045 AACGCCAGCAACCGGCCCTTTTACGGTTCCTGGCCTTTTGTGGCCTTTTGCTCACATG 1986
QY 5413 ttctttcctgcgttatccccctgattctgtggataaccgtattaccgccttttgagtgagct 5472
Db 1985 TTCTTTCCCTGGCTTATCCCTGTATCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCT 1926
QY 5473 gataccgctcgcgcagccgaacgacgagcgcagcagtcagtcagtcagtcagtcagtcagtcag 5532
Db 1925 GATACCGCTCGCCGCAGCCGAACGACGAGCGCAGCGAGTCAGTCAGTCAGTCAGTCAGTCAG 1866
QY 5533 gagcgcctaatacgcgaacccctctccgcgcgttgcccgatttcattaatgcagctgg 5592
Db 1865 GAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTGTGGCCGATTTCATTAATGACGCTGG 1806
QY 5593 cagcacaggtttcccgactggaaagcgggcagtcgagcgaacgcgaataatgtgagttag 5652
Db 1805 CACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAG 1746
QY 5653 ctcaactcattaggcaccgccagggctttacactttatgcttccggctcgtagttgtgga 5712
Db 1745 CTCACCTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGA 1686
QY 5713 attgtgagcggataacaatttcacacaggaacagctatgaccatgattacgccca 5767
|||||

Db 1685 ATTGTGAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGATTACGCCA 1631
RESULT 8
US-08-232-016-23/c
; Sequence 23, Application US/08232016
; Patent No. 5952547
; GENERAL INFORMATION:
; APPLICANT: CORNELISSEN, Marc
; APPLICANT: SOETAERT, Piet
; APPLICANT: STAM, Maïke
; APPLICANT: DOCKX, Jan
; APPLICANT: VAN AARSEN, Roel
; TITLE OF INVENTION: MODIFIED GENES AND THEIR EXPRESSION IN
; TITLE OF INVENTION: PLANT CELLS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,016
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91402920.2
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92400820.4
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K
; REGISTRATION NUMBER: P39,300
; REFERENCE/DOCKET NUMBER: 010830-049
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: plasmid DNA designated as pPS0212
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1785
; OTHER INFORMATION: /note= "Coding region of a
; OTHER INFORMATION: truncated modified bt2 (cryIab) gene, also designated as t
; OTHER INFORMATION: cryIab6 gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1793..2026
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA gene 7."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2396..2921
; OTHER INFORMATION: /note= "35S promoter sequence
; OTHER INFORMATION: derived from Cauliflower mosaic virus."
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: 2922..3581
OTHER INFORMATION: /note- "Coding sequence of
OTHER INFORMATION: chloramphenicol acetyl transferase gene."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3582..4407
OTHER INFORMATION: /note- "3' regulatory sequence
OTHER INFORMATION: containing the polyadenylation site derived from Agrobacterium
OTHER INFORMATION: T-DNA octopine synthase gene."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5600..6457
OTHER INFORMATION: /note- "Sequence complementary to
Patent No. 5952547
OTHER INFORMATION: the coding sequence of the beta-lactamase gene."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7071..7566
OTHER INFORMATION: /note- "TR1' and TR2 promoter
OTHER INFORMATION: derived from Agrobacterium T-DNA (with modified leader with
OTHER INFORMATION: respect to sequence of pJD884 of SEQ ID NO. 22."
US-08-232-016-23

Query Match 43.8%; Score 2527; DB 2; Length 7566;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2627; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3138 actggcgcgtcggttttacaaacgtcgtgactgggaaaaaccctggcgttacccaaacttaacg 3197
DB 7052 ACTGGCGCGTGGTTTACAAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAACTTAATCG 6993
QY 3198 ccttgccagcacatccccctttcgcagcctgaatggcgaatggcgccctgcagtaacatcgtctcct 3257
DB 6992 CCTTGCAGCACATCCCCCTTTCGCCAGCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCCT 6873
QY 3258 ccttcccaacagttgcgcagcctgaatggcgaatggcgccctgcagtaacatcgtctcct 3317
DB 6932 CCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCCT 6873
QY 3318 tacgcatctgctcggtatttcacacccgcataatggtgcactctcagtaacatcgtctcga 3377
DB 6872 TAGGCATCTGTGCGGTATTTTCACACCCGCATATGTTGCACCTCTCAGTACAATCTGCTCTGA 6813
QY 3378 tgcgcgcatagttaagccagcccgacaccccgcaacaccccgctgacgcgccctgacgggc 3437
DB 6812 TGCCGCATAGTTAAGCCAGCCCGCACACCCGCCAACACCCGCTGACCGCCCTGACGGGC 6753
QY 3438 ttgtctgctcccgccatccgcttacagacaaagctgtgacccgtctccgggagctgcatgtg 3497
DB 6752 TTGTCTGTCTCCGGCATCCGCTTACAGACAAAGCTGTGACCGCTCTCCGGAGCTGCATGTG 6693
QY 3498 tcagagggtttcacccgtcatcacccgaaacgcgcgagacgaaagggcctcgtgatacgcct 3557
DB 6692 TCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCT 6633
QY 3558 attttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcg 3617
DB 6632 ATTTTATAGGTTAATGTATGATATAAATGGTTTCTTAGACGCTCAGTGGCACCTTTTCG 6573
QY 3618 gggaaatgvcgcggaacccctattgtttatttttctaataacattcaaatatgtatcc 3677
DB 6572 GGGAAATGTGCGCGGAACCCCTATTGTGTTATTTTCTAAATACATTCAAATATGTATCC 6513
QY 3678 gctcatgagacaataaccctgatataatgcttcaataatattgaaaaaggaagatgatgag 3737
DB 6512 GCTCATGAGACAATAACCCTGTATAAATGCTTCAATAATATTGAAAGGAAGAGTATGAG 6453
QY 3738 tattcaacattccgtgtcgcccttattcccttttttgcggcattttgccttccctgtttt 3797
DB 6452 TATTCAACATTTCGGTGTGCGCCCTTATTCCTCTTTTTCGGGCAATTTTCCTTCCCTTCC 6393
QY 3798 tgctcacccgaaaacgctggtgaaagtaaaagatgctgaagatcagttggtgacgaggt 3857

DB 6392 TGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT 6333
QY 3858 gggttacatgaactggatctcaacagcggtgaagatcccttgagagttttcgcggcgaaga 3917
DB 6332 GGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCGAAGA 6273
QY 3918 acgttttccaatgatgagcaccttttaagtctcgtatgtggcgcggtattatcccgat 3977
DB 6272 ACGTTTCCAATGATGAGCACCTTTTAAAGTCTGCTATGTGGCGCGGTATATCCCGTAT 6213
QY 3978 tgacgcccgggcaagagcaactcggctgcgcgcacatacactattctcagaatgacttggtga 4037
DB 6212 TGACGCCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTGA 6153
QY 4038 gtactcaccagtcacagaaaagcatcttaccggatggcatgacagtaagagaattatgcag 4097
DB 6152 GTACTCACCACTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATGCAG 6093
QY 4098 tgcgtccataaccatgagtataacactcggcgcaacttacttctgacaacgacgcgag 4157
DB 6092 TGCTGCCATAACCATGAGTGATAACACTGCGCGCAACTTACTTCTGACAAAGATCGGAGG 6033
QY 4158 accgaaggagctaaccgcttttttgcaacaacatgggggatcatgtaactgccttgatcg 4217
DB 6032 ACCGAAGGAGCTAACCGCTTTTGTGCACAACATGGGGGATCATGTAACTCGCCTTGATCG 5973
QY 4218 ttgggaaccggagctgaatgaagccatacccaaacgacgagcgtgacacccagatgcctgt 4277
DB 5972 TTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCATGCTGT 5913
QY 4278 agcaatggcaacaacgcttgccgcaaacatttaactggcgaactacttactctagcttccc 4337
DB 5912 AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTTACTTCTAGCTTCCCG 5853
QY 4338 gcaacaattaaatagactggatggagggcggaataaagtgcaggaccacttctgcgctcggc 4397
DB 5852 GCAACAATTATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCCTTCTGCGCTCGGC 5793
QY 4398 ccttccggctggctggtttattgctgataaactggagccggtgagcgtggttctcggcg 4457
DB 5792 CCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGCTGAGCGTGGTCTCGCGG 5733
QY 4458 tatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgac 4517
DB 5732 TATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC 5673
QY 4518 ggggagtcaggcaactatggatgaacgaaatagacagatcgcgtgagataggtgcctcact 4577
DB 5672 GGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGTGGCTCACT 5613
QY 4578 gattaagcattggtaactgtcagaccaagtttactcatatataacttttagattgattaaa 4637
DB 5612 GATTAAGCATTTGTAACCTGTGACACCAAGTTTACTCATATATATCTTTAGATTGATTAAA 5553
QY 4638 acttcatttttaatttaaaaggatctagggtgaagatcccttttgataatctcatgaccaa 4697
DB 5552 ACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCATGACCAA 5493
QY 4698 aatcccttaacgtgagtttctccactgagcgtcagaccccgtagaaaaagatcaaaag 4757
DB 5492 AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAAG 5433
QY 4758 atcttcttgagatcctttttctgcgcgctaactcgtgcttgcaaacacaaaaaacacc 4817
DB 5432 ATCTTCTTGAGATCCTTTTTCGCGCGTAACTGCTGCTTGCAACAAAAAACCCACC 5373
QY 4818 gctaccagcgggtggtttgtttgcgggatcaagagctaccaactcttttccgaaggtaac 4877
DB 5372 GCTACCAGCGGTGTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC 5313
QY 4878 tggcttcagcagcgcagataccacaaactgtccttctagtgtagccgtagtgccca 4937

Db 5312 TGGCTTCAGCAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGTGGCCA 5253
QY 4938 ccacttcaagaactctgttagcaccgcctacatacctcgtctctgctaatactctgttaccagt 4997
Db 5252 CCACCTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGT 5193
QY 4998 ggctgctgccagtggcgataaagtcgtgtcttaaccggggttggaactcaagaacatagttacc 5057
Db 5192 GGCTGCTGCCAGTGGCGATAAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACC 5133
QY 5058 ggataaggcgagcgggtcgggctgaacgggggttcgtgcacacagccccagcttgagcgc 5117
Db 5132 GGATAAGGCGCAGCGGTGGGTGAAACGGGGGTTCGTGCACACAGCCCGCTTGGAGCG 5073
QY 5118 aacgacctacacgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcc 5177
Db 5072 AAGGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCACGCTTCC 5013
QY 5178 cgaaggagaaaaggcgacaggtatccggtaagcggcagggtcggaacagagagcgcac 5237
Db 5012 CGAAGGGAGAAAGSCGACAGGTATCCGGTAAGCGGCGAGGTTCGGAACAGGAGCGCAC 4953
QY 5238 gagggagcttccagggggaaacgcctggtatctttatagtcctgctgcgggttcgccacct 5297
Db 4952 GAGGGAGCTTCCAGGGGGAACGCCCTGGTATCTTATAGTCTCTGCGGGTTTCGCCACCT 4893
QY 5298 ctgacttgagcgtgatattttgtgatgctcgtcagggggcgagccttatgcaacatgttctt 5357
Db 4892 CTGACTTGAGCGTCGATTTTGTGATGCTCGTCAAGGGGGCGGAGCCTATGGAANAACGC 4833
QY 5358 cagcaacgcggcctttttacggttcctggccttttgccttttgcacatgttctt 5417
Db 4832 CAGCAACGCGGCCCTTTTACGGTTCTCTGGCCTTTTGTGGCCTTTTGTGCTCACATGTTCTT 4773
QY 5418 tcctgcgttatccctgattctgtggataacgctattaccgcctttgagtgagctgatac 5477
Db 4772 TCCTGCGTTATCCCTCGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAC 4713
QY 5478 cgctcgccgcagcgaacgacgcgagcgcagcagtgagcaggaagcggaaagcgc 5537
Db 4712 CGCTCGCCGCGCAGCGCAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAGCG 4653
QY 5538 cccaatacgcgaacccgctctccccgcgcttgccgattcattaatgagctggcacga 5597
Db 4652 CCCAATACGCAACCGCCCTCTCCCCCGCGCTTGGCCGATTCAATTAATGACGCTGGCACGA 4593
QY 5598 caggtttccgcactggaaaagcgggcagtgagcgcgaacgaattaatgtgagttagctcac 5657
Db 4592 CAGGTTTCCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGCTCAC 4533
QY 5658 tcattagggaccccgcttttacactttatgcttccggctcgtatgttgtggaattgt 5717
Db 4532 TCATTAGGCACCCAGCGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAATTGT 4473
QY 5718 gagcggataacaatttcacacaggaacagctatgaccatgattacgcc 5766
Db 4472 GAGCGGATAACAATTTACACACAGGAACAGCTATGACCATGATTACGCC 4424

RESULT 9

US-08-232-016-22/c
; Sequence 22, Application US/08232016
; Patent No. 5952547
; GENERAL INFORMATION:
; APPLICANT: CORNELISSEN, Marc
; APPLICANT: SOETAERT, Piet
; APPLICANT: STAM, Maïke
; APPLICANT: DOCKX, Jan
; APPLICANT: VAN AARSEN, Roel
; TITLE OF INVENTION: MODIFIED GENES AND THEIR EXPRESSION IN
; TITLE OF INVENTION: PLANT CELLS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,016
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91402920.2
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92400820.4
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K
; REGISTRATION NUMBER: P39,300
; REFERENCE/DOCKET NUMBER: 010830-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: plasmid DNA designated as pJD884
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1869
; OTHER INFORMATION: /note= "Coding region of a
; OTHER INFORMATION: truncated bt2 (cryIAb) gene, also designated as the bt884
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1877..2110
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA gene 7."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2480..3005
; OTHER INFORMATION: /note= "35S promoter sequence
; OTHER INFORMATION: derived from Cauliflower mosaic virus."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3006..3665
; OTHER INFORMATION: /note= "Coding sequence of
; OTHER INFORMATION: chloramphenicol acetyl transferase gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3666..4491
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA octopine synthase gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5684..6541
; OTHER INFORMATION: /note= "Sequence complementary to
; Patent No. 5952547
; OTHER INFORMATION: the coding sequence of the beta-lactamase gene."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7155..7639

; OTHER INFORMATION: /note= "TR1 and TR2' promoter ; OTHER INFORMATION: derived from Agrobacterium T-DNA." US-08-232-016-22									
Query Match		43.8%;	Score 2527;	DB 2;	Length 7639;				
Best Local Similarity		99.9%;	Pred. No. 0;						
Matches 2627;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;				
QY	3138	actggccgctgcttttacaacgctgctgactgggaaaaacccctggcgttacccaaacttaacg	3197						
Db	7136	ACTGGCCGTCGTTTACAAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATCG	7077						
QY	3198	ccttgagcacatccccctttccgcagctggcgtaataagcgaagaggccgcaccgatcg	3257						
Db	7076	CCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCG	7017						
QY	3258	cccttcccaaacagttagcgcagcctgaatggcgaaatggcgccctgatcgcggtattttctcct	3317						
Db	7016	CCCTTCCCAACAGTTGGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTCTCCT	6957						
QY	3318	tacgcatctgtgcggtatttcacacccgcataatggtgcactctcagttacaatctgctctga	3377						
Db	6956	TACGCATCTGTGCGGTATTTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGCTCTGA	6897						
QY	3378	tgcgcgcatagttaaagccagccccgcagacacccgcgaacccgcgctgacgcgccctgacgggc	3437						
Db	6896	TGCCGCATAGTTAAGCCAGCCCCGACACCCGCCCAACACCCGCTGACGCCCTGACGGGC	6837						
QY	3438	ttgtctgctcccggtcatccgcttacagacaagctgtgacgctcctccgggagctgcatgtg	3497						
Db	6836	TTGTCTGTCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTG	6777						
QY	3498	tcagagggttttcaccgtcatcacccgaacccgcgcgagacgaaaaggccctcgtgatacgcct	3557						
Db	6776	TCAGAGGTTTTCACCGTCTATCACCGAAACCGCGGAGACGAAAGGGCCTCGTGATACGCC	6717						
QY	3558	atttttatagggttaatgtcatgataataatgggtttcttagacgtcagggtggcacttttcg	3617						
Db	6716	ATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACCTTTTCG	6657						
QY	3618	gggaaatgtgcgcggaacccctatttgtttatttttotaataacattcaaatatgtatcc	3677						
Db	6656	GGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCC	6597						
QY	3678	gctcatgagacaataaccctgataaaatgcttcaataataattgaaaaaggaagatagag	3737						
Db	6596	GCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATAATTGAAAAAGGAAGATAGAG	6537						
QY	3738	tattcaacatttcctggtgcgccttattcccttttttgcggcattttgccttccctgtttt	3797						
Db	6536	TATTCACATTTCCGCTGTGCGCCCTTATTCCTTTTTCGCGCATTTTTCCTTCTGTTTT	6477						
QY	3798	tgtcacccagaaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagt	3857						
Db	6476	TGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT	6417						
QY	3858	gggttacatcgaactggatctcaacacagcggtaagatccttgagagttttcgccccgaaga	3917						
Db	6416	GGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCCGAAGA	6357						
QY	3918	acgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatccccgtat	3977						
Db	6356	ACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCCGTAT	6297						
QY	3978	tgcgcgggggcaagagcaactcggctgcgcgcatacacactattctcagaatgacttgggtga	4037						
Db	6296	TGACGCGCGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAATGACTTGGTTGA	6237						
QY	4038	gtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaattatgcag	4097						
Db	6236	GTACTCACCCAGTCACAGAAAACCATCTTACGGATGGCATGCAGTAAGAGAAATTATGCAG	6177						

QY	4098	tgctgccataaccatgagtataaactgcggaactgcggcaacttacttctgacaacgatcggagg	4157						
Db	6176	TGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGG	6117						
QY	4158	accgaaggagctaaccgccttttttcgacaacatgggggatcatgataactcgccttgatcg	4217						
Db	6116	ACCGAAGGAGCTAACCGCTTTTTCGACAACATGGGGGATCATGTAACTCGCCTTGATCG	6057						
QY	4218	ttgggaaccgggagctgaatgaagccataaccaaaacgacgagcgtagacaccacgatgcctgt	4277						
Db	6056	TTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGT	5997						
QY	4278	agcaatgggcaacaacgttgcgcaactattaaactggcgaaactacttactctagcttcccg	4337						
Db	5996	AGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTTACTTCCCG	5937						
QY	4338	gcaacaattaatagactggatggagcgggataaaagtgcaggaccacttctgcgtcggc	4397						
Db	5936	GCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACTTCTCGCCTCGGC	5877						
QY	4398	ccctccggctggctggtttattgtctgataaaatctggagccgggtgagcgtgggtctcgcgg	4457						
Db	5876	CCTTCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGG	5817						
QY	4458	tatcatgtcagcactggggccagatgggtaagccctcccgtaactcgttagttatctacacgac	4517						
Db	5816	TATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC	5757						
QY	4518	ggggagtcaggcaactatggatgaacgaaaatagacagatcgcctgagataggtgcctcact	4577						
Db	5756	GGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT	5697						
QY	4578	gattaagcattggtaactgtcagaccaaagtttactcatatataacttttagattgatttaa	4637						
Db	5696	GATTAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAA	5637						
QY	4638	acttcatttttaatttaaaggatctaggtgaagatcccttttttgataatctcatgaccaa	4697						
Db	5636	ACTTCATTTTAAATTAAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCATGACCA	5577						
QY	4698	aatcccttaacgtgagtttcttcctcactgagcgtcagaccccgtagaaaaaatcaaaagg	4757						
Db	5576	AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATCAAAG	5517						
QY	4758	atctcttgagatccctttttcttcgcgtaaatctgctgctgtgtaacaaacaaaaaccacc	4817						
Db	5516	ATCTTCTTGAGATCCCTTTTTCGCGGTAATCTGCTGCTTGCAACAAAAAACACCACC	5457						
QY	4818	gctaccagcgggtggtttgtttgcgggatcaagagctaccaactctttttccgaaggtaac	4877						
Db	5456	GCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC	5397						
QY	4878	tggtctcagcagagcgagataccaaaactgtccctctagtgtagccgtagttaggcca	4937						
Db	5396	TGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCA	5337						
QY	4938	ccacttcaagaactctgtagcacgcgcctacataacctcgtctctgctaactcctgttaccagt	4997						
Db	5336	CCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGTCTGTAACTCTGTACCAGT	5277						
QY	4998	ggctgctgccagtgggcgataaagtgtgtcttaccgggttggactcaagacgatagttacc	5057						
Db	5276	GGCTGCTGCCAGTGGCGGATAAGTCGTGTCTTACC GG GTTGACTCAAGACGATAGTTACC	5217						
QY	5058	ggataaggcgagcggctgggtgaacgggggttcgtgcacacagcccagcttggagcg	5117						
Db	5216	GGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGGTTCGTGTCACACAGCCCAGCTTGGAGCG	5157						
QY	5118	aacgacctacaccgaaactgagatcacctacagcgtgagctatgagaaaagcgccacgcttcc	5177						
Db	5156	AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTTGAGAAAAGCGCCACGCTTCC	5097						
QY	5178	cgaaggagaaaaggcgagacaggtatccgggtaagcggcagggttcggaaacagagagcgcac	5237						

Db	5096	CGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGACAGGTCGGAACAGGAGCGCAC	5037
Qy	5238	gagggagcttcagggggaacgcctgtatctttatagtcctgtcggtttcgcacact	5297
Db	5036	GAGGGAGCTTCAGGGGGAACGCCCTGGTATCTTTATAGTCCGTGTGGGTTTCGCCACCT	4977
Qy	5298	ctgacttgagggctgattttttgtgatgctgcgtcaggggggagcgcctatggaaaaacgc	5357
Db	4976	CTGACTTGAGCGTCGATTTTGTGTATGCTGCTCAGGGGGCGGAGCCTATGGAACACGC	4917
Qy	5358	cagcaacggggccttttttacgggttcctggcccttttctgctggcccttttctgcacatgttctt	5417
Db	4916	CAGCAACGGCGCCTTTTACGGTTCCTGGCCTTTTGTGGCCTTTTGCTCACATGTTCTT	4857
Qy	5418	tcctgogttatccccctgattctctgtgataaacctgtattaccgcctttgagtgagctgatac	5477
Db	4856	TCCTGCGTTATCCCCCTGATTCTGTGGATAACCCGTATTACCGCCTTTTGAGTGAGCTGATAC	4797
Qy	5478	cgctgcgcgagccgaaacgacgagcagcgagtcagtcagtcagtcagtcagtcagtcagtcag	5537
Db	4796	CGCTCGCGCAGCCGAAACGACCGAGCGAGCGAGTCAGTGAGCGAGGAAGCGGAGCGG	4737
Qy	5538	cccaatacgaacacccgcctctccccgcggttgccgattcattaatgacgctggcacga	5597
Db	4736	CCCAATACGAAACCGCCTCTCCCCGCGGTTGGCCGATTTCATTAAATGCAGCTGGCACGA	4677
Qy	5598	caggtttccgagactggaagcgggcagtcagtcagtcagtcagtcagtcagtcagtcagtcac	5657
Db	4676	CAGGTTTCCGACTGGAAGCGGSCAGTCAGCGCAACGCAATTAAATGAGTTAGCTCAC	4617
Qy	5658	tcattaggcaccgccaggttttacacatttatgctctccgcctcgtatgtgtggaattgt	5717
Db	4616	TCATTAGGCACCCAGGCTTTACACTTATGCTTCCGGCTCGTATGTTGTGGAATTGT	4557
Qy	5718	gagcgggatacaaatttcacacaggaacagctatgaccatgattacgcc	5766
Db	4556	GAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGATTACGCC	4508
RESULT 10			
US-09-194-285-7			
; Sequence 7, Application US/09194285			
; Patent No. 6355479			
; GENERAL INFORMATION:			
; APPLICANT: Webb, Susan R.			
; APPLICANT: Wingvist, Ola			
; APPLICANT: Karlsson, Lars			
; APPLICANT: Jackson, Michael R.			
; APPLICANT: Peterson, Per A.			
; TITLE OF INVENTION: MHC Class II Antigen Presenting Systems			
; FILE REFERENCE: TSRI 536.1			
; CURRENT APPLICATION NUMBER: US/09/194,285			
; CURRENT FILING DATE: 1999-04-12			
; PRIOR APPLICATION NUMBER: PCT/US97/08697			
; PRIOR FILING DATE: 1997-05-22			
; PRIOR APPLICATION NUMBER: US 60/018,175			
; PRIOR FILING DATE: 1996-05-23			
; NUMBER OF SEQ ID NOS: 56			
; SOFTWARE: FastSeq for Windows version 4.0			
; SEQ ID NO 7			
; LENGTH: 4713			
; TYPE: DNA			
; ORGANISM: Mus musculus			
US-09-194-285-7			
Query Match 43.8%; Score 2526; DB 4; Length 4713;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2626; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	3137	cactggccgtgttttacacgctgactgggaaacccctggcgttaccacaacttaac	3196

Db	2086	cactggccgtgtgttttacacgctgactgggaaacccctggcgttaccacaacttaac	2145
Qy	3197	gccttgagcaacatcccccttttcgccagctggcgtaataagcgaagagggccgcacgcac	3256
Db	2146	gccttgagcaacatcccccttttcgccagctggcgtaataagcgaagagggccgcacgcac	2205
Qy	3257	gcccttcccaacagttgcgcagccctgaatggcgaatggcgcctgatgggtattttctcc	3316
Db	2206	gcccttcccaacagttgcgcagccctgaatggcgaatggcgcctgatgggtattttctcc	2265
Qy	3317	ttacgcacatctgtgcgggtattttcacacccgcataatatgggtgcactctcagtcacaatctgctctg	3376
Db	2266	ttacgcacatctgtgcgggtattttcacacccgcataatatgggtgcactctcagtcacaatctgctctg	2325
Qy	3377	atgcgcgcacatagttaagccagccccgcacacccgcgcacacccgcctgacgcgcctgacggg	3436
Db	2326	atgcgcgcacatagttaagccagccccgcacacccgcgcacacccgcctgacgcgcctgacggg	2385
Qy	3437	cttgtctgtcccgccatccgccttacagacaagctgtgaccgtctccgggagctgcatgt	3496
Db	2386	cttgtctgtcccgccatccgccttacagacaagctgtgaccgtctccgggagctgcatgt	2445
Qy	3497	gtcagaggtttttcacccgtcatcacccgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgc	3556
Db	2446	gtcagaggtttttcacccgtcatcacccgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgc	2505
Qy	3557	tattttttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcacttttc	3616
Db	2506	tattttttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcacttttc	2565
Qy	3617	ggggaaaatgtgcgcgggaacccctattttgtttatttttctaaatacatcacaataatgtatc	3676
Db	2566	ggggaaaatgtgcgcgggaacccctattttgtttatttttctaaatacatcacaataatgtatc	2625
Qy	3677	cgctcatgagacaataaacccctgataaatgcttcaataataattgaaaaaggaagagtatga	3736
Db	2626	cgctcatgagacaataaacccctgataaatgcttcaataataattgaaaaaggaagagtatga	2685
Qy	3737	gtattcaacatttccgtgtcgcgccttattcccttttttgcggcattttgccttccctgttt	3796
Db	2686	gtattcaacatttccgtgtcgcgccttattcccttttttgcggcattttgccttccctgttt	2745
Qy	3797	ttgctcacccagaaaacgcctggtgaaagtataaaagatgctgaagatcagttgggtgcacgag	3856
Db	2746	ttgctcacccagaaaacgcctggtgaaagtataaaagatgctgaagatcagttgggtgcacgag	2805
Qy	3857	tgggttacatcgaaactggatctcaacagcgggtaagatcccttgagagttttcgccccgaag	3916
Db	2806	tgggttacatcgaaactggatctcaacagcgggtaagatcccttgagagttttcgccccgaag	2865
Qy	3917	aacgtttttccaaatgatgagcacatttttaaagttctgctatgtggtggcggtattatcccgta	3976
Db	2866	aacgtttttccaaatgatgagcacatttttaaagttctgctatgtggtggcggtattatcccgta	2925
Qy	3977	ttgacgcggggcaagagcaactcgtgcgcgcatacacactattctcagaatgacttgggtg	4036
Db	2926	ttgacgcggggcaagagcaactcgtgcgcgcatacacactattctcagaatgacttgggtg	2985
Qy	4037	agtactcacagtcacagaaaagcatcttcacggatggcatgacagtaagagaattatgca	4096
Db	2986	agtactcacagtcacagaaaagcatcttcacggatggcatgacagtaagagaattatgca	3045
Qy	4097	gtgctgccataaacatgagtataaacactgcggcccaacttacttctgacaacgatcggag	4156
Db	3046	gtgctgccataaacatgagtataaacactgcggcccaacttacttctgacaacgatcggag	3105
Qy	4157	gaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatc	4216
Db	3106	gaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatc	3165
Qy	4217	gttgggaaacccggagctgaatgaagccataccaaacgcagcgtgacaccacgatgcctg	4276

Db 3166 gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctg 3225
Qy 4277 tagcaatggcaacaacagttgcgcaaaactattaaactggcgaaactactactctagctccc 4336
Db 3226 tagcaatggcaacaacgttgcgcaaaactattaaactggcgaaactactactctagctccc 3285
Qy 4337 ggcaacaattaatagactggatggagcgggataaaagttgcaggaccacttctgcgctcg 4396
Db 3286 ggcaacaattaatagactggatggagcgggataaaagttgcaggaccacttctgcgctcg 3345
Qy 4397 cccttcgggctggctggtttattgctgataaaatctggagccggtgagcgtg9ggtctcg 4456
Db 3346 cccttcgggctggctggtttattgctgataaaatctggagccggtgagcgtg9ggtctcg 3405
Qy 4457 gtatcattgcagcactggggccagatgggtaagccctcccgtatcgttagttatctacacga 4516
Db 3406 gtatcattgcagcactggggccagatgggtaagccctcccgtatcgttagttatctacacga 3465
Qy 4517 cggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcac 4576
Db 3466 cggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcac 3525
Qy 4577 tgattaagcattggttaactgtcagaccaagtttactcatatatacttttagattgatttaa 4636
Db 3526 tgattaagcattggttaactgtcagaccaagtttactcatatatacttttagattgatttaa 3585
Qy 4637 aacttcatttttaatttaaaggatctaggtgaagatcccttttttgataaatctcatgacca 4696
Db 3586 aacttcatttttaatttaaaggatctaggtgaagatcccttttttgataaatctcatgacca 3645
Qy 4697 aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaagatcaaaag 4756
Db 3646 aaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaagatcaaaag 3705
Qy 4757 gatctcttgagatccctttttttctgctgtaactgtcgtgctgctgctgcaaaacaaacccac 4816
Db 3706 gatctcttgagatccctttttttctgctgtaactgtcgtgctgctgcaaaacaaacccac 3765
Qy 4817 cgctaccagcgggtggtttgtttgcgggatcaagagctaccacactctttttccgaaggttaa 4876
Db 3766 cgctaccagcgggtggtttgtttgcgggatcaagagctaccacactctttttccgaaggttaa 3825
Qy 4877 ctggcttcagcagagcgcagatatacaaaatactgtccttctagtgtagcgttagtgcc 4936
Db 3826 ctggcttcagcagagcgcagatatacaaaatactgtccttctagtgtagcgttagtgcc 3885
Qy 4937 accacttcaagaactctgtagcacgcgcctacatacctcgtcgtcgttaactcctgttaccag 4996
Db 3886 accacttcaagaactctgtagcacgcgcctacatacctcgtcgtcgttaactcctgttaccag 3945
Qy 4997 tggctgctgccagtggcgataaagtgtgtcttaccgggttggaactcaagacgatatgtac 5056
Db 3946 tggctgctgccagtggcgataaagtgtgtcttaccgggttggaactcaagacgatatgtac 4005
Qy 5057 cggataaggcgcagcggctgggtgaaacgggggttcgtgcaacacagcccagcttggagc 5116
Db 4006 cggataaggcgcagcggctgggtgaaacgggggttcgtgcaacacagcccagcttggagc 4065
Qy 5117 gaacgacctacaccgaactgagatacctacagcgtgagctatgagaaaagcgcacgcttc 5176
Db 4066 gaacgacctacaccgaactgagatacctacagcgtgagctatgagaaaagcgcacgcttc 4125
Qy 5177 ccgaaggggagaaaaggcggacaggtatccggtaagcggcagggtcggaacagagagcgca 5236
Db 4126 ccgaaggggagaaaaggcggacaggtatccggtaagcggcagggtcggaacagagagcgca 4185
Qy 5237 cgagggagcttccagggggaaaacgcctggtatctttatagtcctgctcgggttccgcccac 5296
Db 4186 cgagggagcttccagggggaaaacgcctggtatctttatagtcctgctcgggttccgcccac 4245
Qy 5297 tctgacttgagcgtcgattttttgatgctcgtaggggggggagcctatggaaaaaacg 5356
Db 4246 tctgacttgagcgtcgattttttgatgctcgtaggggggggagcctatggaaaaaacg 4305

Qy 5357 ccagcaacgcggcccttttttacggttcctgcttcttctgctggccttttgcacatgttct 5416
Db 4306 ccagcaacgcggcccttttttacggttcctgcttcttctgctggccttttgcacatgttct 4365
Qy 5417 ttctgctgttatccctgattctgtggataaccgtattaccgcttctgagtgagctgata 5476
Db 4366 ttctgctgttatccctgattctgtggataaccgtattaccgcttctgagtgagctgata 4425
Qy 5477 ccgctcgcgcagccggaacgacccgagcgcagcagtcagtgagcggaggaagcgaagagc 5536
Db 4426 ccgctcgcgcagccggaacgacccgagcgcagcagtcagtgagcggaggaagcgaagagc 4485
Qy 5537 gcccaatacgcgaacacgcctctctcccgcggttgccgattcattaatgcagctggcacg 5596
Db 4486 gcccaatacgcgaacacgcctctctcccgcggttgccgattcattaatgcagctggcacg 4545
Qy 5597 acaggtttcccgactggaaagcgggcagtcagtcagcgaacgcaattaatgcagttagctca 5656
Db 4546 acaggtttcccgactggaaagcgggcagtcagtcagcgaacgcaattaatgcagttagctca 4605
Qy 5657 ctcataggcaacccaggtttacactttatgcttccggctcgtatgttctggtggaattg 5716
Db 4606 ctcataggcaacccaggtttacactttatgcttccggctcgtatgttctggtggaattg 4665
Qy 5717 tgagcggataacaatttcacacagggaacagctatgacctgattacg 5764
Db 4666 tgagcggataacaatttcacacagggaacagctatgacctgattacg 4713

RESULT 11
US-09-194-285-8
; Sequence 8, Application US/09194285
; Patent No. 6355479
; GENERAL INFORMATION:
; APPLICANT: Webb, Susan R.
; APPLICANT: Wingvist, Ola
; APPLICANT: Karlsson, Lars
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; TITLE OF INVENTION: MHC Class II Antigen Presenting Systems
; TITLE OF INVENTION: and Methods for Activating CD4+ T Cells
; FILE REFERENCE: TSRI 536.1
; CURRENT APPLICATION NUMBER: US/09/194,285
; CURRENT FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: PCT/US97/08697
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 60/018,175
; PRIOR FILING DATE: 1996-05-23
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 4724
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-194-285-8

Query Match 43.8%; Score 2526; DB 4; Length 4724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2626; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3137 cactggccgctgcttttacaaactcgtgactgggaaacccctggcttaccacaaacttaac 3196
Db 2097 cactggccgctgcttttacaaactcgtgactgggaaacccctggcttaccacaaacttaac 2156
Qy 3197 gccttgcagcacatccccctttcgcagctggcgtaatagcgaagggcccgaccgac 3256
Db 2157 gccttgcagcacatccccctttcgcagctggcgtaatagcgaagggcccgaccgac 2216
Qy 3257 gcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatcggtattttctcc 3316
Db 2217 gcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatcggtattttctcc 2276

QY	3317	ttacgcatctgtcgggtattttcacaccgcgcatatggtgcactctcagttacaatctgctctg	3377
Db	2277	ttacgcatctgtcgggtattttcacaccgcgcatatggtgcactctcagttacaatctgctctg	2336
QY	3377	atgccgcatagtttaagccagccccgcagacacccgcgaacaccccgctgacgcgccttgacggg	3436
Db	2337	atgccgcatagtttaagccagccccgcagacacccgcgaacaccccgctgacgcgccttgacggg	2396
QY	3437	cttgctctccccggcatcccgcttacagacaaagtgtgaccggtctccgggagctgcatgt	3496
Db	2397	cttgctctccccggcatcccgcttacagacaaagtgtgaccggtctccgggagctgcatgt	2456
QY	3497	gtcagaggtttttcacccgtcatcacccgaaacgcgcgcgagacgaaaggcctcgtgatacgcc	3556
Db	2457	gtcagaggtttttcacccgtcatcacccgaaacgcgcgcgagacgaaaggcctcgtgatacgcc	2516
QY	3557	tattttataggttaatgtcatgataataatggttttcttaaatacattcgaatgtgacttttc	3616
Db	2517	tattttataggttaatgtcatgataataatggttttcttaaatacattcgaatgtgacttttc	2576
QY	3617	ggggaaatgtcgcgggaacccctatttggttatttttcttaaatacattcgaatgtgacttc	3676
Db	2577	ggggaaatgtcgcgggaacccctatttggttatttttcttaaatacattcgaatgtgacttc	2636
QY	3677	cgctcatgagacaataaacctgatataatgcttcaataatattgaaaaaggaagagtatga	3736
Db	2637	cgctcatgagacaataaacctgatataatgcttcaataatattgaaaaaggaagagtatga	2696
QY	3737	gtattcaacatttccgtgtcgccttatttcccttttttgcggcatttttgccttccctgttt	3796
Db	2697	gtattcaacatttccgtgtcgccttatttcccttttttgcggcatttttgccttccctgttt	2756
QY	3797	ttgctcaccagaaacgctggtgaaagttaaagatgctgaagatcagttgggtgcacgag	3856
Db	2757	ttgctcaccagaaacgctggtgaaagttaaagatgctgaagatcagttgggtgcacgag	2816
QY	3857	tgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaag	3916
Db	2817	tgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaag	2876
QY	3917	aacgttttccaatgatgagcacattttaagttctgcttatgtggcgggtattatcccgta	3976
Db	2877	aacgttttccaatgatgagcacattttaagttctgcttatgtggcgggtattatcccgta	2936
QY	3977	ttgacgcggggcaagagcaactcggtcgcgcgcatacacactattctcagaatgacttggttg	4036
Db	2937	ttgacgcggggcaagagcaactcggtcgcgcgcatacacactattctcagaatgacttggttg	2996
QY	4037	agtactcacagtcacagaaaaagcatcttaccggatggcatgacagtaagagaattatgca	4096
Db	2997	agtactcacagtcacagaaaaagcatcttaccggatggcatgacagtaagagaattatgca	3056
QY	4097	gtgctgccataaacatgagtataacactcggcccaacttacttctgacaacgacgcggag	4156
Db	3057	gtgctgccataaacatgagtataacactcggcccaacttacttctgacaacgacgcggag	3116
QY	4157	gaccgaaggagctaaccgcttttttgacacaacatgggggatcatgtgaactgcgcttgatc	4216
Db	3117	gaccgaaggagctaaccgcttttttgacacaacatgggggatcatgtgaactgcgcttgatc	3176
QY	4217	gttgggaacccgggtgaatgaagccataccaaaacgacgagcgtgacacacgacgtgcctg	4276
Db	3177	gttgggaacccgggtgaatgaagccataccaaaacgacgagcgtgacacacgacgtgcctg	3236
QY	4277	tagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctagcttccc	4336
Db	3237	tagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctagcttccc	3296
QY	4337	ggcaacaattaatagactggatggaggcggaataaaagttagcaggaccacttctgcgctcgg	4396
Db	3297	ggcaacaattaatagactggatggaggcggaataaaagttagcaggaccacttctgcgctcgg	3356

QY	4397	cccttcggtgctgggtttattgctgataaaatctgagccggtgagcgtgggtctgcgcg	4455
Db	3357	cccttcggtgctgggtttattgctgataaaatctgagccggtgagcgtgggtctgcgcg	3416
QY	4457	gtatcattgcagcactggggccagatggttaagccctccgtatcgtagttatctacacga	4516
Db	3417	gtatcattgcagcactggggccagatggttaagccctccgtatcgtagttatctacacga	3476
QY	4517	cggggagtcaggccaactatggatgaacgaaatagacagatcgctgagataggctgcctcac	4576
Db	3477	cggggagtcaggccaactatggatgaacgaaatagacagatcgctgagataggctgcctcac	3536
QY	4577	tgattaaagcattggtaactgtcagaccaagtttactcatatatacttttagattgatttaa	4636
Db	3537	tgattaaagcattggtaactgtcagaccaagtttactcatatatacttttagattgatttaa	3596
QY	4637	aaacttcatttttaatttaaaaggatctagggtgaagatcctttttgataatctcatgacca	4696
Db	3597	aaacttcatttttaatttaaaaggatctagggtgaagatcctttttgataatctcatgacca	3656
QY	4697	aaatcccttaacgtgagtttctcgttccactgagcgtcagaccccgtagaaaaagatcaaaag	4756
Db	3657	aaatcccttaacgtgagtttctcgttccactgagcgtcagaccccgtagaaaaagatcaaaag	3716
QY	4757	gatcttcttgagatcccttttttctgcgcgtaactctgtgcttgcaaacaaaaaaaccac	4816
Db	3717	gatcttcttgagatcccttttttctgcgcgtaactctgtgcttgcaaacaaaaaaaccac	3776
QY	4817	cgctaccagcgggttggttttgccggatcaagagctaccacactcttttccgaaggtaa	4876
Db	3777	cgctaccagcgggttggttttgccggatcaagagctaccacactcttttccgaaggtaa	3836
QY	4877	ctggcttcagcagagcgcagatataccaaatactgtccttctagtgtagccgtagttaggcc	4936
Db	3837	ctggcttcagcagagcgcagatataccaaatactgtccttctagtgtagccgtagttaggcc	3896
QY	4937	accactcaagaaactctgtagcaccgcctacatacctcgtcctgtctaactcctgttaccag	4996
Db	3897	accactcaagaaactctgtagcaccgcctacatacctcgtcctgtctaactcctgttaccag	3956
QY	4997	tggctgctgccagtggcgataaagtcgtcttaccgggttggaactcaagacgatagttac	5056
Db	3957	tggctgctgccagtggcgataaagtcgtcttaccgggttggaactcaagacgatagttac	4016
QY	5057	cggataaggcgagcgggtcgggtgaacgggggttcgtgcacacagcccagcttggagc	5116
Db	4017	cggataaggcgagcgggtcgggtgaacgggggttcgtgcacacagcccagcttggagc	4076
QY	5117	gaacgacctacacgcgaactgagataacctacagcgtgagcatlgagaaagcgccacgcttc	5176
Db	4077	gaacgacctacacgcgaactgagataacctacagcgtgagcatlgagaaagcgccacgcttc	4136
QY	5177	ccgaaaggagaaaaggcggacaggtatccgggttaagcggggtcggaaacagagagcgca	5236
Db	4137	ccgaaaggagaaaaggcggacaggtatccgggttaagcggggtcggaaacagagagcgca	4196
QY	5237	cgaggagcttccaggggggaaacgcctggatactttatagtcctgtcgggtttcggccacc	5296
Db	4197	cgaggagcttccaggggggaaacgcctggatactttatagtcctgtcgggtttcggccacc	4256
QY	5297	tctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagccctatggaaaaaacg	5356
Db	4257	tctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagccctatggaaaaaacg	4316
QY	5357	ccagcaacgcggcctttttacgggttccttgcccttttgctggccttttgctcacatgttct	5416
Db	4317	ccagcaacgcggcctttttacgggttccttgcccttttgctggccttttgctcacatgttct	4376
QY	5417	tctctgcgttatccctgatctctgtggataaaccgattaccgccttttgagtgagctgata	5476
Db	4377	tctctgcgttatccctgatctctgtggataaaccgattaccgccttttgagtgagctgata	4436
QY	5477	ccgctcgcgcgagccgaacgaccgcgagcgcagcagtcagtcagtcgaggaagcggaaagacg	5536

Db 4437 cgcgtcgccgcagccgaacgcagcgagcgagtcagtgagcgaggaagcggagagc 4496
QY 5537 gcccaatacgcgaacccgcctctcccccgcgcgttgcccgattcattaatgcagctggcacg 5596
Db 4497 gcccaatacgcgaacccgcctctcccccgcgcgttgcccgattcattaatgcagctggcacg 4556
QY 5597 acaggtttcccgactggaagcggcgagtgagcgcgaacgcgaattaatgtgagttagctca 5656
Db 4557 acaggtttcccgactggaagcggcgagtgagcgcgaacgcgaattaatgtgagttagctca 4616
QY 5657 ctcaattagccacccaggctttacactttatgcttccggctcgtatgttgtgtggaattg 5716
Db 4617 ctcaattagccacccaggctttacactttatgcttccggctcgtatgttgtgtggaattg 4676
QY 5717 tgagcggataacaatttcacacaggaaacagctatgaccatgattacg 5764
Db 4677 tgagcggataacaatttcacacaggaaacagctatgaccatgattacg 4724

RESULT 12
US-09-238-356-28/c
; Sequence 28, Application US/09238356
; Patent No. 6312683
; GENERAL INFORMATION:
; APPLICANT: Kingsman, et al
; TITLE OF INVENTION: Retroviral Vectors
; FILE REFERENCE: 674523-2006
; CURRENT APPLICATION NUMBER: US/09/238,356
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: PCT/GB/03876
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28
; LENGTH: 7387
; TYPE: DNA
; ORGANISM: Artificial Sequence, primer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(7387)
; OTHER INFORMATION: plasmid vector
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AX003207
; DATABASE ENTRY DATE: 2000-08-24
; RELEVANT RESIDUES: (1)..(7387)
US-09-238-356-28

Query Match 43.0%; Score 2482; DB 4; Length 7387;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2632; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3133 aattcactggccgtcgttttacacgcgtcgtgactggggaaaacccctggcgttacccaactt 3192
Db 7356 AATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTT 7297
QY 3193 aatcgcccttgccagcacatccccctttcgccagctggcgtaataagcgaagagcccgacc 3252
Db 7296 AATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGGCCCGCCACC 7237
QY 3253 gatcgcccttcccaacagttggcgcagcctgaatggcgcaatggcgccctgatgcggtattt 3312
Db 7236 GATCGCCCTTCCCAACAGTTGGCGAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTT 7177
QY 3313 ctccctacgcactctgtgcggtatttcacacccgcataatggtgcactctcagtaacaatcgc 3372
Db 7176 CTCCCTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGC 7117
QY 3373 tctgatgccgcatagttaagccagccccgcacacccgcgaacaccccgctgacgcgccctga 3432
Db 7116 TCTGATGCCGCATAGTTAAGCCAGCCCCGCACACCCCGCCACACCCCGCTGACGCCGCCCTGA 7057

QY 3433 cgggcttgtctgctccccggcatcgcttacagacaagctgtgaccgctctccgggagctgc 3492
Db 7056 CGGCTTGTTGTTCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC 6997
QY 3493 atgtgtcagaggtttttcacccgtcatcacggaaacgcgcgagagcaaaaggcctcgtgata 3552
Db 6996 ATGTCTCAGAGGTTTTCACCGTCAACCGAAACGCGCGAGAGCAAGGGCCTCGTGATA 6937
QY 3553 cgcctattttataggttaatgtcatgataataatggtttcttagacgctcaggtggcact 3612
Db 6936 CGCCTATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACT 6877
QY 3613 tttcggggaaaatgtgcgcggaaacccctatttgtttattttctaaatacatctcaaatatg 3672
Db 6876 TTTCCGGGAAAATGTCCGGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATATG 6817
QY 3673 tatcgctcatgagacaataaacctgataaaatgcttcaataatattgaaaaaggaagagt 3732
Db 6816 TATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGT 6757
QY 3733 atgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttct 3792
Db 6756 ATGAGTATTCAACATTTCCGTTGTCGCCCTTATTCCCTTTTTCGGGCATTTTGCCCTTCT 6697
QY 3793 gttttgtccaccagaaaacgctggtgaaagtataaagatgctgaagatcagttgggtgca 3852
Db 6696 GTTTTGTCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 6637
QY 3853 cgagtgggttacatcgaaactggaatctcaacacagcggtaagatccttgagagtttgcgcc 3912
Db 6636 CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTTCGCC 6577
QY 3913 gaagaacgttttccaaatgatgagcaacttttaaagttctgtatgtggcggtattatcc 3972
Db 6576 GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTATTATCC 6517
QY 3973 cgtattgacgcgggcaagagcaactcggcgccgcatacactattctcagaatgacttg 4032
Db 6516 CGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTG 6457
QY 4033 gttgagtactcaccagtcacagaaaaagcatcttacggatggcgatgacagtaagagaatta 4092
Db 6456 GTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTA 6397
QY 4093 tgcagtgctgccataaaccatgagtgataaacactgcggccaacttacttctgacaaacgatc 4152
Db 6396 TGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAACGATC 6337
QY 4153 ggagaccggaaggagctaaccgcttttttgacaaacatgggggatcatgtaactcgcctt 4212
Db 6336 GGAGGACCGGAAGGAGCTAACCCGCTTTTTCACAACATGGGGGATCATGTAACCTCGCCTT 6277
QY 4213 gatcgttgggaaccggagctgaatgaagccataccaaaacgagcgtgacacacagatg 4272
Db 6276 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATG 6217
QY 4273 cctgtagcaaatggcaacaacgtttgcgcaaaactattaaactggcgaaactacttacttagct 4332
Db 6216 CCTGTAGCAATGCCAACACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCT 6157
QY 4333 tccccggcaacaattaatagactggatggagcggaataaagtgcaggaccacttctgcgc 4392
Db 6156 TCCCGSCAACAAATTAATAGACTGGATGGAGGGCGGATAAAGTTGCAGGACCACCTTCTGCGC 6097
QY 4393 tcggcccttccggctggctgtttattgtgtataaaatctggagccggtgagcgtgggtct 4452
Db 6096 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCGGTGAGCGTGGGTCT 6037
QY 4453 cgcggtatcatgtcagcactggggccagatggtaagccctcccgctatcgtagtattctac 4512
Db 6036 CGCGGTATCATTGCAGCACTGGGGCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC 5977
QY 4513 acgacggggagtcaggcaactatggatgaacgaataagacagatcgctgagataggtgcc 4572

Db 61 TCGCCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGA 120
QY 3255 tcgccccttcccaacagttgcgagccctgaatggcgaaatggcgccctgatgcggtattttct 3314
Db 121 TCGCCCTTCCCAACAGTTGCCGAGCCCTGAATGGCGAATGGCGCCTGATGGGGTATTTTCT 180
QY 3315 ccttacgcattctgtgcggtattttcacaccgcgcatatggtgcactctcaagtacaactctgctc 3374
Db 181 CCTTACGCATCTGTGCGGTATTTTACACCCGCATATGGTGCACCTCTCAGTACAATCTGCTC 240
QY 3375 tgatgccgcgatagttaagccagccccgcacacccgcgaacaccccgctgacgcgccctgacg 3434
Db 241 TGATGCCGATAGTTAAGCCAGCCCCGACACCCGCCAACACCCCGCTGACGCGCCCTGACG 300
QY 3435 ggcttgctcgtcccgccatccgcttacagacaagctgtgacgcgtctccgggagctgcat 3494
Db 301 GGCTTGCTCCTCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCTGCAT 360
QY 3495 gtgtcagaggtttttcaccgtcatcacccgaaacgcgcgagacgaaaggccctcgtgatacg 3554
Db 361 GTGTCAGAGGTTTTTACCCGTCATCACCGAAGCGCGGAGACGAAAGGGCCCTCGTGATACG 420
QY 3555 cctatttttataaggttaatgtcatgataataatggtttcttagacgtcaggtggcactttt 3614
Db 421 CCTATTTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTT 480
QY 3615 tcggggaaatgctgcggaacccctatttgtttatttttctaaatacattcaaatatgta 3674
Db 481 TCGGGGAAATGTGCGCGGAACCCCTATTGTATTATTTTCTAAATACATTCAAAATATGTA 540
QY 3675 tccgctcatagacaataaacccctgataaatgcttcaataatatgaaaaaggaagatat 3734
Db 541 TCCGCTCATGAGACAAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATAT 600
QY 3735 gagtattcaacatttcggtgcgcccctattcccttttttgcggcattttgccttcctgt 3794
Db 601 GAGTATTCACATTTCCGTGTCGCCCTTATTCCTTTTTCGGGCATTTTGCCCTTCCTGT 660
QY 3795 ttttgcctcaccagaaacgcctggtgaaagtaaaagatgctgaagatcagttgggtgcacg 3854
Db 661 TTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACG 720
QY 3855 agtgggttacatcgaactggatctcaacagcggtaagatcccttgagagttttcgccccga 3914
Db 721 AGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGA 780
QY 3915 agaacttttccaatgatgagcacttttaaagtctcgtatgtggcgcggtattatcccg 3974
Db 781 AGAACGTTTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGGGTATTATCCCG 840
QY 3975 tattgacgcgggcaagagcaactcggtcgcccgcatacacactattctcagaatgacttgg 4034
Db 841 TATTGACGCCCGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGT 900
QY 4035 tgagtactcaccagtcacagaaaaagcatcttacggatggcatgacatgacagtaagagaattatg 4094
Db 901 TGAGTACTCACCAGTCAAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATG 960
QY 4095 cagtgctgcataaaccatgagtgtataaactcggccaaacttacttctgacaacgatcgg 4154
Db 961 CAGTGTGCCATAACCATGAGTGATAAACACTGCGGCCAACTTACTTCTGACAACGATCCG 1020
QY 4155 aggaccgaaggagctaaccgctttttgcacaacatggggatcatgtaaactcgccctga 4214
Db 1021 AGGACCGAAGGAGCTAACCGCTTTTTTGACAACATGGGGATCATCTAACTCGCCTTGA 1080
QY 4215 tcggtgggaacgggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcc 4274
Db 1081 TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACCCAGTGCC 1140
QY 4275 tgtagcaatggcaacaacgcttgcgcaaacatttaactggcggaacttacttactctagcttc 4334

Db 1141 TGAGCAATGGCAACAACGTTTGGCGAAACTATTAACTGGCGAACTACTTACTCTAGCTTC 1200
QY 4335 ccggcaacaattaatagactggatggaggcggaataaagttgcaggaccacttctgcgctc 4394
Db 1201 CCGGCAACAATTAATAGACTGGATGGAGCGCGATAAAGTTGCAGGACCCTTCTGCGCTC 1260
QY 4395 ggcccttccggctggctggtttatttgctgataaaatctggagccggtgagcgtgggtctcg 4454
Db 1261 GGCCCTTCCGGCTGGCTGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCG 1320
QY 4455 cggatatcattgcagcactggggccagatggttaagccctcccgtatcgtagttatctacac 4514
Db 1321 CGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC 1380
QY 4515 gacggggagtcaggcaactatggtgaacgaataagacagatcgtgagataggtgcctc 4574
Db 1381 GACGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGTAGGTGCCTC 1440
QY 4575 actgattaagcattggttaactgtcagaccgaagttaactatatatactttagattgattt 4634
Db 1441 ACTGATTAAGCATTTGGTAACTGTGAGACCAAGTTTACTCATATATACTTTAGATTGATT 1500
QY 4635 aaaacttcatattttaatttaaaggatctaggtgaagatccctttttgataaatctcatgac 4694
Db 1501 AAAACTTCATTTTAAATTTAAAGGATCTAGSTGAAGATCCTTTTGTGATAATCTCATGAC 1560
QY 4695 caaaaatcccttaacgctgagttttcgtccactgagcgtcagacccccgtagaaaaagatcaa 4754
Db 1561 CAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAA 1620
QY 4755 aggatcttcttgagatcccttttttctgcggttaactcgtcgtcttgcaacaaaaaac 4814
Db 1621 AGGATCTTCTGAGATCCTTTTTTCTGCGGTAATCTGCTGTTGCAACAAAAAAC 1680
QY 4815 accgctaccagcgggtgtgttgcgggatcaagagctaccaactcttttccgaaggt 4874
Db 1681 ACCGCTACCAGCGGTGTTGTTTGC CGGATCAAGAGCTACCAACTCTTTTCCGAAGGT 1740
QY 4875 aactggcttcagcagagcgcagataccacaaatactgtccttctagtgtagccgtagttagg 4934
Db 1741 AACTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTTAGTGTAGCCGTAGTTAGG 1800
QY 4935 ccaccacttcaagaactctgtagcaccgcctacatacctcgtctcgtctaatcctgtttacc 4994
Db 1801 CCACCACCTTCAAGAACTCTGTAGCACCCGCCCTACATACTCTGCTCTGTCTGTATTACC 1860
QY 4995 agtggctgctgcagtgccgataaagctgctgtcttaccgggttgactcaagacgatagtt 5054
Db 1861 AGTGGCTGCTGCCAGTGGCGATAAAGTCTGTCTTACC GGGTTGGACTCAAGACGATAGTT 1920
QY 5055 accggataaaggcgagcggctcgggctgaacggggggttcgtgcacacagccagccttggga 5114
Db 1921 ACCGGATAAGGCGAGCGGTGCGGCTGAACGGGGGTTCTGTGCACACAGCCAGCTTGA 1980
QY 5115 gcgaacgacctacacccgaactgagataacctacagcgtgagctatgagaaaagccacgct 5174
Db 1981 GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTTGAGAAAAGGCCACGCT 2040
QY 5175 tcccgaaggaggagaaaaggcggaacaggtatccgggtaagcggcagggtcggaaacaggagagcg 5234
Db 2041 TCCCGAAGGGAGAAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCG 2100
QY 5235 cagcagggagcttccagggggaaacgcctggtattctttatagtcctcgtcggggtttcgcca 5294
Db 2101 CACGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCTCTGTCGGGTTTCGCCA 2160
QY 5295 cctctgacttgagcgtcgaatttttgtgatgctcgtcagggggcgagccttatggaaaaa 5354
Db 2161 CCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAA 2220
QY 5355 cgccagcaacgcggccttttttacgggttcctggccttttgcggccttttgcctcacatgtt 5414
Db 2221 CGCCAGCAACGGCGCCTTTTTACGGTTCCCTGGCCTTTTGTGGCCTTTTGTGCTCACATGTT 2280

QY 5415 ctttcctgcgttatccctgattctgtggataaacgctattaccgcctttgagtgagctga 5474
|||||
Db 2281 CTTTCCTGCGTTATCCCTTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA 2340

QY 5475 taccgctgcgcgcagcgaacgaccgagcgcagcagtgagtcagtgagcgaaggaaga 5534
|||||
Db 2341 TACCGCTCGCGCGACGCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGA 2400

QY 5535 gcgcaccaatacgcgaacccgcctctccccgcgcgttggccgattcatttaataatgcagctggca 5594
|||||
Db 2401 GCGCCCAATACGCAAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCA 2460

QY 5595 cgacaggtttcccgactggaaagcgggcagtgagcgaacgcgaattaataatgtagcttagct 5654
|||||
Db 2461 CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCT 2520

QY 5655 cactcattaggcaccacccaggcctttacacctttatgcttccggctcgtatgttgtggaat 5714
|||||
Db 2521 CACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAAT 2580

QY 5715 tgtgagcggataaacaatttcacacaggaacagctatgaccatgattacgcga 5767
|||||
Db 2581 TGTGAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCCA 2633

RESULT 14
US-08-844-274-20
; Sequence 20, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; EARLIER FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-08-844-274-20

Query Match 42.9%; Score 2475; DB 4; Length 7560;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 2625; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3137 cactggccgtcgttttacaaactcgtgactgggaaaaacccctggcgttacccaacttaatc 3196
|||||
Db 4766 cactggccgtcgttttacaaactcgtgactgggaaaaacccctggcgttacccaacttaatc 4825

QY 3197 gccttgacgacatcccccttcgccagctggcgaatggcgccctgatgcggtattttctcc 3256
|||||
Db 4826 gccttgacgacatcccccttcgccagctggcgaatggcgccctgatgcggtattttctcc 4885

QY 3257 gcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatgcggtattttctcc 3316
|||||
Db 4886 gcccttcccaacagttgcgcagcctgaatggcgaatggcgccctgatgcggtattttctcc 4945

QY 3317 ttacgcatctgtgggtattttcacaccgcataatggtgcactctcagtagacaatctgctctg 3376
|||||
Db 4946 ttacgcatctgtgggtattttcacaccgcataatggtgcactctcagtagacaatctgctctg 5005

QY 3377 atgcgcgcatagttaagccagccccgcagacacccgccaaacacccgcgtgacgcgccttgacgag 3436
|||||
Db 5006 atgcgcgcatagttaagccagccccgcagacacccgccaaacacccgcgtgacgcgccttgacgag 5065

QY 3437 cttgtctgtctcccgcatcccgcttacagacaagctgtgaccgctctccgggagctgcatgt 3496
|||||
Db 5066 cttgtctgtctcccgcatcccgcttacagacaagctgtgaccgctctccgggagctgcatgt 5125

QY 3497 gtcagaggttttcaccgctcatcacccgaaacgcgcgagacgaaaggcctcgtgatacgcc 3556
|||||
Db 5126 gtcagaggttttcaccgctcatcacccgaaacgcgcgagacgaaaggcctcgtgatacgcc 5185

QY 3557 tattttataggttaatgtcatgataataatggttttcttagacgtcaggtggcacttttc 3616
|||||
Db 5186 tattttataggttaatgtcatgataataatggttttcttagacgtcaggtggcacttttc 5245

QY 3617 ggggaaatgtgcgcggaaacccctatttgtttatttttctaaatacattcaaatatgtatc 3676
|||||
Db 5246 ggggaaatgtgcgcggaaacccctatttgtttatttttctaaatacattcaaatatgtatc 5305

QY 3677 cgctcatgagacaataacccctgataaaatgcttcaataatatgaaaaaggaagagtatga 3736
|||||
Db 5306 cgctcatgagacaataacccctgataaaatgcttcaataatatgaaaaaggaagagtatga 5365

QY 3737 gtattcaacatttcctgtgcgccttattcccttttttgcggcattttgccttccctgttt 3796
|||||
Db 5366 gtattcaacatttcctgtgcgccttattcccttttttgcggcattttgccttccctgttt 5425

QY 3797 ttgctcaccagaaaacgctggtgaaagtataaagatgctgaagatcagttgggtgcacgag 3856
|||||
Db 5426 ttgctcaccagaaaacgctggtgaaagtataaagatgctgaagatcagttgggtgcacgag 5485

QY 3857 tgggttacatcgaaactgatctcaacagcggtaagatccttgagagttttgcgccgaag 3916
|||||
Db 5486 tgggttacatcgaaactgatctcaacagcggtaagatccttgagagttttgcgccgaag 5545

QY 3917 aacgttttcccaatgatgagcacattttaagttctctgtatgtggcgggtattatcccgta 3976
|||||
Db 5546 aacgttttcccaatgatgagcacattttaagttctctgtatgtggcgggtattatcccgta 5605

QY 3977 ttgacgcggggcaagagcaactcggtcgcgcgcatacactattctcagaatgacttggttg 4036
|||||
Db 5606 ttgacgcggggcaagagcaactcggtcgcgcgcatacactattctcagaatgacttggttg 5665

QY 4037 agtactcaccagtcacagaaaaagcatctttacggatggcagatggcagtaagagaattatgca 4096
|||||
Db 5666 agtactcaccagtcacagaaaaagcatctttacggatggcagatggcagtaagagaattatgca 5725

QY 4097 gtgctgccataaaccatgagtataaacactcggcgcgaacttacttctgacaacacgagag 4156
|||||
Db 5726 gtgctgccataaaccatgagtataaacactcggcgcgaacttacttctgacaacacgagag 5785

QY 4157 gaccgaaggagctaacgcgttttttgacacaacatgggggatcatgtaactcgccttgatc 4216
|||||
Db 5786 gaccgaaggagctaacgcgttttttgacacaacatgggggatcatgtaactcgccttgatc 5845

QY 4217 gttgggaacccggagctgaatgaagccataccaaaacgacgagcgtgacacccacgagtcctg 4276
|||||
Db 5846 gttgggaacccggcgtgaatgaagccataccaaaacgacgagcgtgacacccacgagtcctg 5905

QY 4277 tagcaatggcaacaacgcttgcgcgaactatttaactggcgaactacttactctagcttccc 4336
|||||
Db 5906 tagcaatggcaacaacgcttgcgcgaactatttaactggcgaactacttactctagcttccc 5965

QY 4337 ggcaacaattaatagactggatggagcgggataaagttgcaggaccacttctgcgctcgg 4396
|||||
Db 5966 ggcaacaattaatagactggatggagcgggataaagttgcaggaccacttctgcgctcgg 6025

QY 4397 cccttccggctggtggtttattgctgataaaatctggagccggtgagcgtgggtctcgcg 4456
|||||
Db 6026 cccttccggctggtggtttattgctgataaaatctggagccggtgagcgtgggtctcgcg 6085

QY 4457 gtatcattgcagcactggggccagatggtaagccctcccgatatcgtagttattctacacga 4516

QY 3317 ttacgcatctgtcggttatattcaaccgc--atatgtgcactctcagtacaaatctgctc 3374
|||||
Db 11102 TTACGCATCTGTGCGGTATTTACACCGCATATATGGTGCACCTCTCAGTACAATCTGCTC 11043

QY 3375 tgatgcgcatagtttaagccagcccgacacccgccaaacaccccgctgacgcgccctgacg 3434
|||||
Db 11042 TGATGCCGATAGTTAAGCCAGCCCGACACCCGCCAACACCCCGCTGACGCCCTGACG 10983

QY 3435 ggcttgctgctcccgcatccgcttacagacaaagctgtgaccgctctccgggagctgcat 3494
|||||
Db 10982 GGCCTGTCTGCTCCCGCATCCGCTTACAGACAAAGCTGTGACCGTCTCCGGGAGCTGCAT 10923

QY 3495 gtgtcagagggttttccacgcgtcatcacccgaaacgcgcgagacgaaaggcctcgtgatacg 3554
|||||
Db 10922 GTGTGAGAGGTTTTCACCGTCTACCGGAACCGCGGAGACGAAAGGGCCTCGTGATACG 10863

QY 3555 cctatttttataggttaagtcatgataataatgggtttctttagacgcgtcaggtggcacctt 3614
|||||
Db 10862 CCTATTTTATAGGTTAATGTCTATGATAATAATGGTTCTTAGACGTCAGGTGGCACTTT 10803

QY 3615 tcggggaaatgtgcggaacccctatttgtttatttttctaatacatccaatatgta 3674
|||||
Db 10802 TCGGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATATGTA 10743

QY 3675 tccgctcatgagacaataaacctgataaaatgcttcaataatatgaaaaaggaagatgat 3734
|||||
Db 10742 TCCGCTCATGAGACAATAACCCCTGATAAAATGCTTCAATAATAATTGAAAAAGGAAGATAT 10683

QY 3735 gagtattcaacatttccggtgtcgcccttattcccttttttggcgcattttgccttccgt 3794
|||||
Db 10682 GAGTATTCAACATTTCCGTTGTCGCCCTTATTCCCTTTTTCGCGCATTTTGCCCTTCCGTG 10623

QY 3795 ttttgcctacccagaaaacgctggtgaaaagtaaaagatgctgaagatcagttgggtgcacg 3854
|||||
Db 10622 TTTTGTCTACCCAGAAAACGCTGGTGTAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACG 10563

QY 3855 agtgggttacatcgaaactggatctcaacagcggtaagatcccttgagagtttctgcgccga 3914
|||||
Db 10562 AGTGGGTACATCGAACTGGAATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGA 10503

QY 3915 agaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccg 3974
|||||
Db 10502 AGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCG 10443

QY 3975 tattgacgcccgggcaagagcaactgggtcgccgcatacactattctcagaatgacttgg 4034
|||||
Db 10442 TATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGT 10383

QY 4035 tgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaaattatg 4094
|||||
Db 10382 TGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATG 10323

QY 4095 cagtgtgcgcataaaccatgagtataaacactgcggccaaacttaacttctgacaaacgatcgg 4154
|||||
Db 10322 CAGTGTGCCATAAACCATGAGTGATAAACACTGCGGCCAACTTACTTCTGACAAACGATCGG 10263

QY 4155 aggaccgaaggagctaaccgcttttttgacaacatgggggatcaatgtaactcgccttga 4214
|||||
Db 10262 AGGACCGAAGGAGCTAACCCGCTTTTGTGCACAACATGGGGGATCATGTAACCTCGCCTTGA 10203

QY 4215 tcgttgggaacccggagctgaatgaagccataaccaaaacgacgagcgtgacacccagatgcc 4274
|||||
Db 10202 TCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACCACGATGCC 10143

QY 4275 tgtagcaatggcaacaacgttgcgaaactattaaactggcgaaactacttactctagcttc 4334
|||||
Db 10142 TGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGCGGAACCTACTTACTCTAGCTTC 10083

QY 4335 ccggcaacaattaatagactggatggagcggataaaagttgcaggaccacttctcgctc 4394
|||||
Db 10082 CCGGCAACAATTAAATAGACTGGATGGAGCGCGGATAAAGTTGCAAGGACCACTTCTCGCGCTC 10023

QY 4395 ggccttcogcgctggctggtttatttctgctgataaaatctggagccggtgagcgtgggtctcg 4454
|||||
Db 10022 GGCCCTTCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGCTGAGCGTGGTCTCG 9963

QY 4455 cggtatcatgacgacactggggccagatgggtaagccctcccgatccgtagttatctacac 4514
|||||
Db 9962 CGGTATCATTCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC 9903

QY 4515 gacggggagtcaggcaactatggatgaacgaaaatagacagatcgcgtgagataggtgcctc 4574
|||||
Db 9902 GACGGGGAGTCAGGCAACTATGGATGAAGCAAAATAGACAGATCGCTGAGATAGGTGCCTC 9843

QY 4575 actgattaaagcattggttaactgtcagaccaagtcttactcatatatactatttagattt 4634
|||||
Db 9842 ACTGATTAAACATTTGTAACCTGTAGACCAAGTTTACTCATATATACTTTAGATTGATTT 9783

QY 4635 aaaacttcatttttaaatttaaaggatctaggtgaagatcccttttttgataaatctcatgac 4694
|||||
Db 9782 AAAACTTCATTTTAAATTAAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCATGAC 9723

QY 4695 caaaatcccttaacgtgagtttttctgctccactgagcgtcagaccccgtagaaaaagatcaa 4754
|||||
Db 9722 CAAAATCCCTTAACGTGAGTTTTCGTTCACACTGAGCGTCAGACCCCGTAGAAAAGATCAA 9663

QY 4755 aggatcttcttgagatcccttttctgctgcgttaactctgctgcttaactccttttccgaaggt 4814
|||||
Db 9662 AGGATCTTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACC 9603

QY 4815 accgctaccagcgggtgttgtttgcccggatcaagagctaccaaactccttttccgaaggt 4874
|||||
Db 9602 ACCGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGT 9543

QY 4875 aactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttagg 4934
|||||
Db 9542 AACTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTAGG 9483

QY 4935 ccaccacttcaagaactctgtagcacgcgcctacatacctcgcctctgctaaatcctgttacc 4994
|||||
Db 9482 CCACCACCTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACC 9423

QY 4995 agtggctgctgccagtggcgataaagtcgtgcttaccgggttggaactcaagacgatatgtt 5054
|||||
Db 9422 AGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACC GGGTTGGACTCAAGACGATAGTT 9363

QY 5055 accgga taaggcgagcggctgggctgaacgggggggttcgtgcacacagcccagcttgg 5114
|||||
Db 9362 ACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGGTTCGTGACACAGCCCCAGCTTGA 9303

QY 5115 gcgaacgacctacacccaactgagataacctacacgctgagctagagctatgagaaagcgccacgct 5174
|||||
Db 9302 GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCCACGCT 9243

QY 5175 tccccgaaggagaaaggcgagacaggtatcccggttaagcggcagggtcggaaacaggagagcg 5234
|||||
Db 9242 TCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGACG 9183

QY 5235 cagcagggagcttccagggggaaacgcctgggtatctttatagtcctgctcggggttctgcca 5294
|||||
Db 9182 CACGAGGAGCTTCCAGGGGGAACCGCTGGTATCTTTATAGTCTCTGTCGGGTTTCGCCA 9123

QY 5295 cctctgacttgagcgtcgattttttacggttccctggccttttctgctggccttttgcacatgtt 5354
|||||
Db 9122 CCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGAAAAA 9063

QY 5355 cgccagaacgcggcctttttacggttccctggccttttctgctggccttttgcacatgtt 5414
|||||
Db 9062 CGCCAGCAACGCGCGCTTTTACGGTTCTCTGGCCCTTTTGTGGCCCTTTTGTCTCACATGTT 9003

QY 5415 ctttctcggttatccccctgattcttgataaacctgattaccgccttttgagtgagctga 5474
|||||
Db 9002 CTTTCTCGGTTATCCCCCTGATTCTGTGGATAACCGTATTACC GCTTTGAGTGAGCTGA 8943

QY 5475 taccgctgcgcgcagccgaacgacccgagcgcagcgagtcagtgagcgagggaagcgggaaga 5534


```

Db 8942 TACCGCTGCGCGCAGCCGAACGCGAGCGCAGGAGTCAGTGAGCGAGGAAGCGGAAGA 8883
Qy 5535 gcgcccaatacgcacaaaccgcctctcccccgcgcgttgcccgattcaatgcagctggca 5594
Db 8882 GCGCCCAATACGCAACCGCCTCTCCCGCGCGTTGGCCGATTCATTATGCAGCTGGCA 8823
Qy 5595 cgacagggtttcccgactggaaaagcggggcagtgagcgcaacgcaattaatgtgagttagct 5654
Db 8822 CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCT 8763
Qy 5655 cactcaataggcacccccaggctttacacttttatgcttccggctcgatatgtgtggaat 5714
Db 8762 CACTCATTAGGCACCCCGAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAAT 8703
Qy 5715 tgtgagcgggataacaatttcacacaggaacacagctatgacccatgattacg 5764
Db 8702 TGTGAGCGGATAACAATTTTCACACAGGAAACAGCTATGACCATGATTACG 8653
```

Search completed: August 31, 2002, 22:43:50
Job time: 19694 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:17:17 ; Search time 824.08 Seconds
(without alignments)
12015.151 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcattgcctgcaggtc.....ctatgaccatgattacgcc 5767

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_032802:*
1:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5767	100.0	5767	24	Plasmid pTM034. S
2	2777.6	48.2	5897	20	Plasmid pTGP190-1
3	2637.8	45.7	8854	20	DNA encoding chime
4	2637	45.7	3858	22	Plasmid pRK50 used
5	2637	45.7	4960	22	pCMV-I-Cre-pA vect
6	2637	45.7	7332	22	pCMV-I-beta-pA vec
C 7	2637	45.7	8062	22	Plasmid pRK76 used
C 8	2637	45.7	8153	22	Plasmid pRK74 used
9	2635.4	45.7	5711	22	C31-Int expression

10	2633.2	45.7	9335	20	AAV63734	Plasmid PIGP/367 c
11	2632.8	45.7	6898	20	AAV63742	Plasmid UGP232-4 c
12	2631.8	45.6	3753	24	ABA04129	Plasmid PHCE19T(II
13	2631.8	45.6	3755	24	ABA04130	Plasmid PHCE19(II)
C 14	2631.8	45.6	4283	18	AAT86449	DNA encoding hGH,
C 15	2631.8	45.6	4283	19	AAT86610	Epidermal growth f
C 16	2631.8	45.6	4283	21	AA63237	Plasmid pWRG1630 n
C 17	2631.8	45.6	6565	22	AAF55126	Nucleotide sequenc
C 18	2631.8	45.6	6971	22	AAF55124	Nucleotide sequenc
C 19	2631.8	45.6	7558	22	AAF55125	Nucleotide sequenc
C 20	2631.8	45.6	7969	22	AAF55123	Nucleotide sequenc
C 21	2630	45.6	6295	19	AAV02043	Plasmid pWRG3196 e
C 22	2630	45.6	7164	22	AAD10238	Commercial plasmid
C 23	2630	45.6	7383	22	AAS00153	Matrix metalloprot
24	2628.2	45.6	4045	15	AAQ70942	Plasmid pND211 con
25	2628.2	45.6	5848	21	AAC55481	Destination vector
C 26	2628	45.6	4950	18	AAV03801	Retroviral vector
C 27	2628	45.6	4950	22	AAC82936	Transdominant effe
28	2628	45.6	5733	21	AAA95416	Aspergillus nidula
C 29	2628	45.6	9164	20	AA82259	Beta-domain delete
C 30	2628	45.6	9632	20	AA82259	Baculovirus transf
C 31	2628	45.6	9780	20	AA82817	Plasmid pFLAG:BAC.
C 32	2628	45.6	11846	20	AA82261	Factor VIII protei
C 33	2628	45.6	12022	20	AA82260	Factor VIII protei
C 34	2627	45.6	7566	14	AAQ42160	Plasmid pPS0212 co
C 35	2627	45.6	7639	14	AAQ42159	Plasmid pJD884 con
C 36	2626.8	45.5	4776	20	AA877617	Expression constru
C 37	2626.8	45.5	4776	20	AA877614	Expression constru
C 38	2626.4	45.5	4163	13	AAQ32349	Template clone Fab
39	2624.8	45.5	4713	19	AAV12067	Murine IAD alpha c
40	2624.8	45.5	4724	19	AAV12068	Murine IAD beta ch
C 41	2624.8	45.5	5421	21	AAC68299	SV40/APPA plasmid
C 42	2624.8	45.5	5465	20	AA220088	Plasmid PUBIAC. C
C 43	2624.8	45.5	6116	21	AAC68297	Plasmid PUBIAC. C
C 44	2624.8	45.5	6708	21	AAC68295	R15/APPA plasmid c
C 45	2624.8	45.5	9534	18	AAT62072	Vector RP3224E2 en

ALIGNMENTS

RESULT 1

AA517547	
ID	AA517547 standard; DNA; 5767 BP.
XX	
AC	AA517547;
XX	
DT	25-FEB-2002 (first entry)
XX	
DE	Plasmid pTM034.
XX	
KW	Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW	transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW	muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW	organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW	circular; ds.
XX	
OS	Synthetic.
XX	
PN	WO200171014-A2.
XX	
PD	27-SEP-2001.
XX	
PF	16-MAR-2001; 2001WO-US08468.
XX	
PR	17-MAR-2000; 2000US-190440P.
XX	
PA	(MORT/) MOR T.
PA	(SORE/) SOREQ H.
PA	(ARNT/) ARNTZEN C.
PA	(MASO/) MASON H.
PA	(BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX	

PI Mor T, Soreq H, Arntzen C, Mason H;

XX WPI; 2002-055120/07.

XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -

XX
PS Claim 11; Page 28-31; 42pp; English.

XX The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigmine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable
CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pRM034, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.

XX
SQ Sequence 5767 BP; 1390 A; 1521 C; 1495 G; 1361 T; 0 other;

Query Match 100.0%; Score 5767; DB 24; Length 5767;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	agcttgcatgcctgcagggtcaacatggtggagcagcacactctcgtctactccaagaata	60
Db	1	agcttgcatgcctgcagggtcaacatggtggagcagcacactctcgtctactccaagaata	60
QY	61	tcaaaagatacagttcagaagaccagagggtattgagacttttcaacaaagggtaatat	120
Db	61	tcaaaagatacagttcagaagaccagagggtattgagacttttcaacaaagggtaatat	120
QY	121	cgggaaacctcctcgtgattccattgccagctatctgtcacttcacgcgaaaggacagtag	180
Db	121	cgggaaacctcctcgtgattccattgccagctatctgtcacttcacgcgaaaggacagtag	180
QY	181	aaaaggaagatggcttctacaaaatgccatcattgcgataaaggaaaggctatcgttcaag	240
Db	181	aaaaggaagatggcttctacaaaatgccatcattgcgataaaggaaaggctatcgttcaag	240
QY	241	aatgcctctaccgcagatgggtcccaaagatggacccccaccacgcgaggaaacatcgtggaa	300
Db	241	aatgcctctaccgcagatgggtcccaaagatggacccccaccacgcgaggaaacatcgtggaa	300
QY	301	aaagaagacgttccaaccacgtcttcaaaagatggattgattgataacttttcaaca	360
Db	301	aaagaagacgttccaaccacgtcttcaaaagatggattgattgataacttttcaaca	360
QY	361	aagggtaatatcgggaaacctcctcggattccaattgccagctatctgtcacttcacga	420
Db	361	aagggtaatatcgggaaacctcctcggattccaattgccagctatctgtcacttcacga	420
QY	421	aaggacagtagaaaaggaaagatggcttctacaaaatgccatcattgcgataaagggaaggc	480
Db	421	aaggacagtagaaaaggaaagatggcttctacaaaatgccatcattgcgataaagggaaggc	480

QY	481	tatcgttcaagaatgcctctaccgacagtgttccaaaagatggacccccaccacgagga	540
Db	481	tatcgttcaagaatgcctctaccgacagtgttccaaaagatggacccccaccacgagga	540
QY	541	acatcgtggaaaaaagaagacgttccaaccacgtcttcaaaagcaagtggatgattgata	600
Db	541	acatcgtggaaaaaagaagacgttccaaccacgtcttcaaaagcaagtggatgattgata	600
QY	601	tctccactgacgtaagggtgacgcacacaatcccactatccttcgcgaagacccttcctcta	660
Db	601	tctccactgacgtaagggtgacgcacacaatcccactatccttcgcgaagacccttcctcta	660
QY	661	tataagggaagtcatctcatttgagaggacctgcagaataattctcaacacaacatat	720
Db	661	tataagggaagtcatctcatttgagaggacctgcagaataattctcaacacaacatat	720
QY	721	acaaacaaacgaatctcaagcaatcaagcattctacttctattgcagcaatttaaatca	780
Db	721	acaaacaaacgaatctcaagcaatcaagcattctacttctattgcagcaatttaaatca	780
QY	781	tttcttttaaagcaaaagcaattttctgaaaattttcacaccatttacgaacgatatccatg	840
Db	781	tttcttttaaagcaaaagcaattttctgaaaattttcacaccatttacgaacgatatccatg	840
QY	841	gctccccgcagtgctgtgcacacgccttcctcctggcttccccactccttcctcctc	900
Db	841	gctccccgcagtgctgtgcacacgccttcctcctggcttccccactccttcctcctc	900
QY	901	ctctggctcctcgtggaggagtgggggctgagggccggaggatgcagagctgctggtg	960
Db	901	ctctggctcctcgtggaggagtgggggctgagggccggaggatgcagagctgctggtg	960
QY	961	acggtgctggtgggcccggctgcggggcattcgcctgaagacccccggggccctgtctct	1020
Db	961	acggtgctggtgggcccggctgcggggcattcgcctgaagacccccggggccctgtctct	1020
QY	1021	gctttcctgggcatcccccttgcggagccaccatgggacccccgcgtcttctgccaccg	1080
Db	1021	gctttcctgggcatcccccttgcggagccaccatgggacccccgcgtcttctgccaccg	1080
QY	1081	gagcccaagcagccttggtcaggggtggtagacgtacacacttccagagtgtctgtctac	1140
Db	1081	gagcccaagcagccttggtcaggggtggtagacgtacacacttccagagtgtctgtctac	1140
QY	1141	caatatgtggacacccctatacccagggttttgagggcaccgcagatgtggaacccccaccgt	1200
Db	1141	caatatgtggacacccctatacccagggttttgagggcaccgcagatgtggaacccccaccgt	1200
QY	1201	gagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccccgcctacatcc	1260
Db	1201	gagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccccgcctacatcc	1260
QY	1261	cccacccctgtcctcgtctggtatctatgggggtggcttctacagtggggcctcctcttg	1320
Db	1261	cccacccctgtcctcgtctggtatctatgggggtggcttctacagtggggcctcctcttg	1320
QY	1321	gacgtgtacgatggcgcgttcttggtacagggccgagaggaactgctggtgtccatgaac	1380
Db	1321	gacgtgtacgatggcgcgttcttggtacagggccgagaggaactgctggtgtccatgaac	1380
QY	1381	taccgggtgggagcctttggcttccctggccctgcggggagccgagagggccccgggcaat	1440
Db	1381	taccgggtgggagcctttggcttccctggccctgcggggagccgagagggccccgggcaat	1440
QY	1441	gtgggtcctcctggatcagaggctggccctgcagtggggtgcaggagaacgtggcagccttc	1500
Db	1441	gtgggtcctcctggatcagaggctggccctgcagtggggtgcaggagaacgtggcagccttc	1500
QY	1501	gggggtgaccccgacatcagtgacgctgttttggggagagcgcgggagccgctcgttgggc	1560
Db	1501	gggggtgaccccgacatcagtgacgctgttttggggagagcgcgggagccgctcgttgggc	1560

QY

1561

atgcacctgctgtccccgcccagccgggcccctgttccacagggccgtgctgcagagcgggt

1620

Db

1561

atgcacctgctgtccccgcccagccgggcccctgttccacagggccgtgctgcagagcgggt

1620

QY

1621

gcccccaatggacccttgggcccacggttgggcatggagagggcccgtcgcagggccacgcag

1680

Db

1621

gcccccaatggacccttgggcccacggttgggcatggagagggcccgtcgcagggccacgcag

1680

QY

1681

ctggccccaccttgttgggctgtcctccagggcggaactggttgggaatgacacagagctggta

1740

Db

1681

ctggccccaccttgttgggctgtcctccagggcggaactggttgggaatgacacagagctggta

1740

QY

1741

gcctgcccctcgggacacgacgaccagcgaggtcccttgcctgttgaaccacgaaatgggcacgtgctgcct

1800

Db

1741

gcctgcccctcgggacacgacgaccagcgaggtcccttgcctgttgaaccacgaaatgggcacgtgctgcct

1800

QY

1801

caagaaaagcgtcttccggttctccttgcctgtggtagatggagacttccctcagtgac

1860

Db

1801

caagaaaagcgtcttccggttctccttgcctgtggtagatggagacttccctcagtgac

1860

QY

1861

accccagaggccctcatcaacgcgggggagacttccacggcctgcaggtgctggtgtg

1920

Db

1861

accccagaggccctcatcaacgcgggggagacttccacggcctgcaggtgctggtgtg

1920

QY

1921

gtgaaggatgagggctcgtattttctgtgtttacggggcccccaggcttcagcaaaagacaac

1980

Db

1921

gtgaaggatgagggctcgtattttctgtgtttacggggcccccaggcttcagcaaaagacaac

1980

QY

1981

gagtctctcatcagccgggcccaggttcttgcctggccggggtgcgggtcccccaggta

2040

Db

1981

gagtctctcatcagccgggcccaggttcttgcctggccggggtgcgggtcccccaggta

2040

QY

2041

agtgacctggcagccgagggctgtggttcttgcattacacagactggctgcacatcccaggac

2100

Db

2041

agtgacctggcagccgagggctgtggttcttgcattacacagactggctgcacatcccaggac

2100

QY

2101

ccggcacgcctgagggagggcccctgagcgatgtggtgggcgaccacaatgtcgtgtgcccc

2160

Db

2101

ccggcacgcctgagggagggcccctgagcgatgtggtgggcgaccacaatgtcgtgtgcccc

2160

QY

2161

gtggcccagctggtggcgactggctgcacaggggtgcccggggtctacgcctacgtcttt

2220

Db

2161

gtggcccagctggtggcgactggctgcacaggggtgcccggggtctacgcctacgtcttt

2220

QY

2221

gaacacccgtgttccacgcgtctccttgcctggcccctgtggaatgggggtgcccccaggtacgag

2280

Db

2221

gaacacccgtgttccacgcgtctccttgcctggcccctgtggaatgggggtgcccccaggtacgag

2280

QY

2281

atcagagtcatcttcttgggatcccccttggaacccctctcgaaaactacacggcgagagagaaa

2340

Db

2281

atcagagtcatcttcttgggatcccccttggaacccctctcgaaaactacacggcgagagagaaa

2340

QY

2341

atcttcgcccagcgactgatgcgatacttgggccaaacttgcgccgcaaggggatcccaat

2400

Db

2341

atcttcgcccagcgactgatgcgatacttgggccaaacttgcgccgcaaggggatcccaat

2400

QY

2401

gagccccgagaccccaaggccccacaatggcccccgtaacacggcggggctcagcagtac

2460

Db

2401

gagccccgagaccccaaggccccacaatggcccccgtaacacggcggggctcagcagtac

2460

QY

2461

gttagtctgacacctgcggccgctggaggtgcggcggggtgcgcgccagccctgcgcc

2520

Db

2461

gttagtctgacacctgcggccgctggaggtgcggcggggtgcgcgccagccctgcgcc

2520

QY

2521

ttcttggaaacgcgttccctccccaaaattgtctacgcgtacacctgataggtaccgagctctctc

2580

Db

2521

ttcttggaaacgcgttccctccccaaaattgtctacgcgtacacctgataggtaccgagctctctc

2580

QY

2581

aacaaatctagctagagtttgcctcctatctatatgttaataagggtatgctgatatgcactat

2640

Db

2581

aacaaatctagctagagtttgcctcctatctatatgttaataagggtatgctgatatgcactat

2640

QY

2641

tcaaataggagcattagctatgtttgtttaatgtcacttttatgttattgtgggtaagtcacc

2700

Db

2641

tcaaataggagcattagctatgtttgtttaatgtcacttttatgttattgtgggtaagtcacc

2700

QY

2701

taagacactccacgtacacctacgttgttgtctctttaccggctttaataaaatcttctgcct

2760

Db

2701

taagacactccacgtacacctacgttgttgtctctttaccggctttaataaaatcttctgcct

2760

QY

2761

tgttcccatatttactaattatccctttcttcaactaaaaaagaaaattgttataatagtat

2820

Db

2761

tgttcccatatttactaattatccctttcttcaactaaaaaagaaaattgttataatagtat

2820

QY

2821

tagtcttttagaaacatatagaggctctttaattgggtaggttttacaaaattaaactaatataaa

2880

Db

2821

tagtcttttagaaacatatagaggctctttaattgggtaggttttacaaaattaaactaatataaa

2880

QY

2881

atgtcataaaaatccacgctgggttaaaacaaaatgcagaaaaatcgacgtcgtctatttggaccga

2940

Db

2881

atgtcataaaaatccacgctgggttaaaacaaaatgcagaaaaatcgacgtcgtctatttggaccga

2940

QY

2941

cagttgctatttaataataatggggccaccatagtagactgacaaaataaaattacctgacaaca

3000

Db

2941

cagttgctatttaataataatggggccaccatagtagactgacaaaataaaattacctgacaaca

3000

QY

3001

tcgtttcactaaataacaaaacacaaaaggagtgcattttccaggggcattttttgaata

3060

Db

3001

tcgtttcactaaataacaaaacacaaaaggagtgcattttccaggggcattttttgaata

3060

QY

3061

aaaaacagttaaaaggagtgcaatagaaaatatagggggtgtggaaaatagtatttgagca

3120

Db

3061

aaaaacagttaaaaggagtgcaatagaaaatatagggggtgtggaaaatagtatttgagca

3120

QY

3121

cgctcttgaagcgaaatttcactggccgtcgttttacaaacgctcgtgactgggaaaaacccctggc

3180

Db

3121

cgctcttgaagcgaaatttcactggccgtcgttttacaaacgctcgtgactgggaaaaacccctggc

3180

QY

3181

gttaccccaacttaataatcgcccttgcagcacatcccccttttcgccagctggcgtaatagcgaa

3240

Db

3181

gttaccccaacttaataatcgcccttgcagcacatcccccttttcgccagctggcgtaatagcgaa

3240

QY

3241

gaggccccgacccgatcgcccttcccaacagtttgcgcagcctgaatggcggaatggcgccctg

3300

Db

3241

gaggccccgacccgatcgcccttcccaacagtttgcgcagcctgaatggcggaatggcgccctg

3300

QY

3301

atgcgggtattttccttacctagcatctgtgcggtatttccacccgcataatggtgcactctc

3360

Db

3301

atgcgggtattttccttacctagcatctgtgcggtatttccacccgcataatggtgcactctc

3360

QY

3361

agtacaaatctgctctgatccgcatagtttaagccagccccgcacacccgcacacacccgct

3420

Db

3361

agtacaaatctgctctgatccgcatagtttaagccagccccgcacacccgcacacacccgct

3420

QY

3421

gacgcgccctgacgggcttgtctgctccccgcacccgttacagacaaagctgtgaccgctc

3480

Db

3421

gacgcgccctgacgggcttgtctgctccccgcacccgttacagacaaagctgtgaccgctc

3480

QY

3481

tccggggagctgcattgtgtcagaggttttcaacgcgtcatcacccgaaacgcgcgagacgaaaag

3540

Db

3481

tccggggagctgcattgtgtcagaggttttcaacgcgtcatcacccgaaacgcgcgagacgaaaag

3540

QY

3541

ggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacg

3600

Db

3541

ggcctcgtgatacgcctatttttataggttaatgtcatgataataatggtttcttagacg

3600

QY

3601

tcagggtggcacttttcggggaatgtgcgcggaacccctatttgtttatttttctaaata

3660

Db

3601

tcagggtggcacttttcggggaatgtgcgcggaacccctatttgtttatttttctaaata

3660

QY

3661

cattcaaaatatgtatccgcctcatgagacaaataaacctgataaaatgcttcaataatattga

3720

Db

3661

cattcaaaatatgtatccgcctcatgagacaaataaacctgataaaatgcttcaataatattga

3720

QY

3721

aaaagggaagagtatgagtatccaacatttccggtgcgcccttatcccttttttgcgggca

3780

Db

3721

aaaagggaagagtatgagtatccaacatttccggtgcgcccttatcccttttttgcgggca

3780

Db 3721 aaaaggaagatagtagtattcaacattttccgtgctgcgcccttattcccttttttgcggca 3780
QY 3781 ttttgcccttcctgtttttgctcaccagaaaacgctgggtgaaaagtaaaaagatgctgaagat 3840
Db 3781 ttttgcccttcctgtttttgctcaccagaaaacgctgggtgaaaagtaaaaagatgctgaagat 3840
QY 3841 cagttgggtgcacagagtggtttacatcgaactcgatctcaacagcggtaagatcccttgag 3900
Db 3841 cagttgggtgcacagagtggtttacatcgaactcgatctcaacagcggtaagatcccttgag 3900
QY 3901 agttttcgccccgaagaaacgtttttccaatgatgagcacattttaaaagttctgctatgtggc 3960
Db 3901 agttttcgccccgaagaaacgtttttccaatgatgagcacattttaaaagttctgctatgtggc 3960
QY 3961 gcggtattatccccgtattgacgcggggaagcaactcggtcgcgcgcatacactattct 4020
Db 3961 gcggtattatccccgtattgacgcggggaagcaactcggtcgcgcgcatacactattct 4020
QY 4021 cagaatgacttgggtgagtgactcaccagtcacagaaaaagcatcttacggatggcatgaca 4080
Db 4021 cagaatgacttgggtgagtgactcaccagtcacagaaaaagcatcttacggatggcatgaca 4080
QY 4081 gtaagagaaattatgcagtgctgcccataaacatgagtgataaacactgcggccaaacttactt 4140
Db 4081 gtaagagaaattatgcagtgctgcccataaacatgagtgataaacactgcggccaaacttactt 4140
QY 4141 ctgacaaacgatcggagggaacggaaggagctaaacgcgttttttgacaaaacttggggatcat 4200
Db 4141 ctgacaaacgatcggagggaacggaaggagctaaacgcgttttttgacaaaacttggggatcat 4200
QY 4201 gtaactcgcccttgatcgttgggaacccggagctgaatgaagccatacacaacgacgagcgt 4260
Db 4201 gtaactcgcccttgatcgttgggaacccggagctgaatgaagccatacacaacgacgagcgt 4260
QY 4261 gacaccacgatgcctgttagcaattggcaacaacgctgcgcaaaactattaaactggcgaaacta 4320
Db 4261 gacaccacgatgcctgttagcaattggcaacaacgctgcgcaaaactattaaactggcgaaacta 4320
QY 4321 cttactcttagcttcccggaacaataatgactggatggaggcgggataaaagtgcaggga 4380
Db 4321 cttactcttagcttcccggaacaataatgactggatggaggcgggataaaagtgcaggga 4380
QY 4381 ccacttctgcgctcggcccttccggctcggctggtttattgctgataaaactggagccgggt 4440
Db 4381 ccacttctgcgctcggcccttccggctcggctggtttattgctgataaaactggagccgggt 4440
QY 4441 gagcgtgggtctcgcgggtatcattgcagcacctggggccagatggtgaagccctcccgatc 4500
Db 4441 gagcgtgggtctcgcgggtatcattgcagcacctggggccagatggtgaagccctcccgatc 4500
QY 4501 gtagtattctacacgacgagggtgagtcaggcaactatggatgaacgaaatagacagatcgct 4560
Db 4501 gtagtattctacacgacgagggtgagtcaggcaactatggatgaacgaaatagacagatcgct 4560
QY 4561 gagataggtgcctcactgattaaagcattggaactgtcagaccaagtttactcatatata 4620
Db 4561 gagataggtgcctcactgattaaagcattggaactgtcagaccaagtttactcatatata 4620
QY 4621 ctttagattgatttaaaaacttcatttttaatttaaaaggatctaggtgaagatcccttttt 4680
Db 4621 ctttagattgatttaaaaacttcatttttaatttaaaaggatctaggtgaagatcccttttt 4680
QY 4681 gataatctcatgaccaaaaatccccctaaacgtgagtttctgcctccactgagcgtcagacccc 4740
Db 4681 gataatctcatgaccaaaaatccccctaaacgtgagtttctgcctccactgagcgtcagacccc 4740
QY 4741 gtagaaaaagatcaaaaggatctcttgagatcccttttttctgcgcgtaactctgctgttg 4800
Db 4741 gtagaaaaagatcaaaaggatctcttgagatcccttttttctgcgcgtaactctgctgttg 4800
QY 4801 caaacaacaaaaaaccccgctaccagcgggtggtttgtttgcccggatcaagagctaccaact 4860
Db 4801 caaacaacaaaaaaccccgctaccagcgggtggtttgtttgcccggatcaagagctaccaact 4860

RESULT 2
AAV63741
ID AAV63741 standard; DNA; 5897 BP.
XX
AC AAV63741;
XX
DT 12-APR-1999 (first entry)

XX DE Plasmid pTGP190-1 containing 35R promoter and per5 3'UTR.

XX KW Peroxidase; per5 gene; maize; corn; transgenic plant; promoter;

KW root; vector; plasmid pTGP190-1; CamV; ds; circular; cyclic.

XX OS Chimeric - Escherichia coli.

OS Chimeric - Zea mays.

OS Chimeric - maize streak virus.

OS Chimeric - Agrobacterium tumefaciens.

OS Chimeric - Cauliflower mosaic virus.

OS Chimeric - synthetic.

XX FH Key Location/Qualifiers

FT misc_feature 18..20

FT /tag= a

FT /note= "linker"

FT 31..282

FT /tag= b

FT /note= "CamV MCASTRAS nt 7093-7344"

FT 283..290

FT /tag= c

FT /note= "linker"

FT 291..637

FT /tag= d

FT /note= "CamV MCASTRAS nt 7093-7439"

FT 638..657

FT /tag= e

FT /note= "linker"

FT 651..1024

FT /tag= f

FT /note= "maize stream virus leader and Adh1 intron"

FT 658..677

FT /tag= g

FT /note= "MSV nt 167-186"

FT 678..767

FT /tag= h

FT /note= "MSV nt 188-277"

FT 769..978

FT /tag= i

FT /note= "Adh1.5 intron"

FT 979..988

FT /tag= j

FT /note= "linker"

FT 989..1028

FT /tag= k

FT /note= "MSV nt 278-317"

FT 1026..2834

FT /tag= l

FT /product= beta-glucuronidase

FT 2835..2890

FT /tag= m

FT /note= "sequence from pKA882"

FT 2896..3261

FT /tag= n

FT /note= "per5 3'UTR nt 6066-6431"

FT 3268..5897

FT /tag= o

FT /note= "pUC19 sequences"

XX PN W09856921-A1.

XX 17-DEC-1998.

PD 10-JUN-1998; 98WO-US11921.

XX 12-JUN-1997; 97US-0049752.

XX (DOWC) DOW AGROSCIENCES LLC.

XX Ainley M, Armstrong K, Belmar S, Folkerts O, Hopkins N;

PI Menke MA, Paredy D, Petolino JF, Smith K, Woosley A;

XX

DR WPI; 1999-080904/07.

XX New isolated regulatory sequences for transgenic plants - which are

PT derived from the maize root preferential cationic peroxidase protein

PT (per5) gene

XX Example 20; Page 123-126; 150pp; English.

XX This is the nucleotide sequence of pTGP190-1, a plasmid comprising

CC a gene cassette in which the following components are operably

CC joined: the 35T promoter, the GUS gene and the maize root

CC preferential cationic peroxidase per5 gene (see AAV63717) 3'

CC untranslated region (3'UTR). The 35T promoter contains tandem

CC copies of the CamV 35S gene promoter, a deleted version of the Adh1

CC intron 1 and the untranslated leader from the maize streak virus

CC coat protein fused to the beta-glucuronidase gene, and is followed

CC by the nos 3'UTR. Testing in maize callus indicated that the per5

CC 3'UTR functioned as well as nos when the GUS gene was driven by the

CC 35T promoter. The invention relates to isolated regulatory

CC sequences, especially promoter, intron and 3'UTR sequences, of the

CC maize per5 gene. Claimed recombinant gene cassettes comprising

CC per5 regulatory sequences are used to control expression of

CC recombinant genes in selected tissue, especially the root, of

CC transformed plants, particularly maize.

XX Sequence 5897 BP; 1506 A; 1431 C; 1510 G; 1450 T; 0 other;

SQ

Query Match 48.2%; Score 2777.6; DB 20; Length 5897;

Best Local Similarity 69.9%; Pred. No. 0;

Matches 4107; Conservative 0; Mismatches 1564; Indels 201; Gaps 19;

QY 94 ttgagacttttcaacaaagggttaatatcgggaaacctcctcggtattccatgccagcta 153

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 29 tggagacttttcaacaaagggttaatatcgggaaacctcctcggtattccatgccagcta 88

QY 154 tctgtcacttcacgaaaggacagtagaagaaggaagatggcttctacaaaatgccatcatt 213

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 89 tctgtcactttattgtgaagatagtggaagaaggaggtggctcctcacaatgccatcatt 148

QY 214 gcgataaaaggaaaggctatcgttcaagaatgcctctaccgacagtggtcccaagaatgga 273

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 149 gcgataaaaggaaaggccatcgttgaag-atgcctctgccgacagtggtcccaagaatgga 207

QY 274 cccccaccacgagggaacatcgctggaaaaagaagacggttccaaccacgcttccaagaacaa 333

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 208 cccccaccacgaggagcgcctgctggaaaaagaagacggttccaaccacgcttccaagaacaa 267

QY 334 gtggattgatgtgat-----aaactttcaacaaagggttaatatcgggaaacctcc 383

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 268 gtggattgatgtgatcatcgatggagacttttcaacaaagggttaatatccggaacctcc 327

QY 384 tcggattccattgccagctatctgtcacttcacatcgaaaggacagtagaagaaggaagatg 443

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 328 tcggattccattgccagctatctgtcactttattgtgaagatagtggaagaaggaaggtg 387

QY 444 gcttctacaaatgccatcattgcgataaaaggaaaggctatcgttcaagaatgcctctacc 503

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 388 gctcctacaaatgccatcattgcgataaaaggaaaggccatcgttgaag-atgcctctgcc 446

QY 504 gacagtgtgtcccaagaatggacccccaccacgagggaacatcgctggaaaaagaagacgtt 563

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 447 gacagtgtgtcccaagaatggacccccaccacgaggagagcgcctggtggaagaagacgtt 506

QY 564 ccaaccacgtcttcaaaagcaagtggattgatgtgatctccactgacgttaagggatgac 623

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 507 ccaaccacgtcttcaaaagcaagtggattgatgtgatctccactgacgttaagggatgac 566

QY 624 gcacaatccactatccttcgcaagaccccttcctctatatataagggaagtccattcattg 683

Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 567 gcacaatccactatccttcgcaagaccccttcctctatatataagggaagtccattcattg 626

QY 684 gagaggacctcgagaattaattctcaacacacatatatacaaaaacaaacgaatctcaagca 743

||||| 5006 actggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgttaggc 5065

Qy 4936 caccacttcaagaactctgtagcaccgcctacatacctcgtctgttaactcctgttacca 4995

Db 5066 caccacttcaagaactctgtagcaccgcctacatacctcgtctgttaactcctgttacca 5125

Qy 4996 gtggctgctgcagtgccgataaagtcgtgtcttaccgggttggaactcaagacgatatgta 5055

Db 5126 gtggctgctgcagtgccgataaagtcgtgtcttaccgggttggaactcaagacgatatgta 5185

Qy 5056 ccggataaaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggag 5115

Db 5186 ccggataaaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggag 5245

Qy 5116 cgaacgacctacacccgaactgagataacctacacgctgagctatgagaaaagcgcacgctt 5175

Db 5246 cgaacgacctacacccgaactgagataacctacacgctgagcattgagaaagcgcacgctt 5305

Qy 5176 cccgaaggagaaaggcgcagaggtatccggttaagcggcaggttcggaacagggagcgc 5235

Db 5306 cccgaaggagaaaggcgcagaggtatccggttaagcggcaggttcggaacagggagcgc 5365

Qy 5236 acgaggggagctccagggggaaacgcctgggtatctttatagtcctctcgggttcgccac 5295

Db 5366 acgaggggagctccagggggaaacgcctgggtatctttatagtcctctcgggttcgccac 5425

Qy 5296 ctctgacttgagcgtcgatttttgtgagtcctcgtcagggggcgagcctatggaaaaac 5355

Db 5426 ctctgacttgagcgtcgatttttgtgagtcctcgtcagggggcgagcctatggaaaaac 5485

Qy 5356 gccagcaacgcggcctttttacgggttcctggcccttttgcggccttttgcacatgttc 5415

Db 5486 gccagcaacgcggcctttttacgggttcctggcccttttgcggccttttgcacatgttc 5545

Qy 5416 tttcctgcgttatccccctgattctgtggataaacgctattaccgcctttgagtgagctgat 5475

Db 5546 tttcctgcgttatccccctgattctgtggataaacgctattaccgcctttgagtgagctgat 5605

Qy 5476 accgctcgcgcagccgaacgcgagccgagcagtcagtcagtcagtcagtcagtcagtcag 5535

Db 5606 accgctcgcgcagccgaacgcgagccgagcagtcagtcagtcagtcagtcagtcagtcag 5665

Qy 5536 cgcccaatacgcacaaacgcctctcccgcgcttgccgattcattaatgcagctggcac 5595

Db 5666 cgcccaatacgcacaaacgcctctcccgcgcttgccgattcattaatgcagctggcac 5725

Qy 5596 gacaggtttccgcgactggaaagcggcagtgagcgcacgcacgcaatgaatgagttagctc 5655

Db 5726 gacaggtttccgcgactggaaagcggcagtgagcgcacgcacgcaatgaatgagttagctc 5785

Qy 5656 actcattaggcaccacccaggctttacactttatgcttccggctcgtatgttgtgtggaatt 5715

Db 5786 actcattaggcaccacccaggctttacactttatgcttccggctcgtatgttgtgtggaatt 5845

Qy 5716 gtgagcggataacaatttcacacaggaacacagctatgacctatgattacgccca 5767

Db 5846 gtgagcggataacaatttcacacaggaacacagctatgacctatgattacgccca 5897

RESULT 3

AAZ20086

ID AAZ20086 standard; DNA; 8854 BP.

XX

AC AAZ20086;

XX

DT 05-JAN-2000 (first entry)

XX

DE DNA encoding chimeric cry1B insecticide HyFLIB.

XX

KW Insecticide; HyFLIB; crystal protein; delta-endotoxin; toxin;

KW cry1B; cry1A(b); maize; transgenic plant; European corn borer;

KW Ostrinia nubilalis; entomocide; crop protection; biological control;

KW pCIB5520; ss.

XX

OS Bacillus thuringiensis.

OS Synthetic.

XX

FH Key

FT CDS

FT Location/Qualifiers

FT 11..3694

FT /*tag= a

FT /product= "HyFLIB"

XX

PN WO9950293-A1.

XX

PD 07-OCT-1999.

XX

PF 30-MAR-1999; 99WO-EP02175.

XX

PR 01-APR-1998; 98US-0053549.

XX

PA (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PI Desai NM;

XX

DR WPI; 1999-601323/51.

DR P-PSDB; AAY31990.

XX

PT Chimeric insecticidal protein comprising cryIB and cryIA(b) portions

XX

PS Claim 3; Page 34-42; 85pp; English.

XX

CC This is the nucleotide sequence of a synthetic gene that codes for

CC a chimeric insecticidal protein, termed hyFLIB (see AAY31990), of 1227

CC amino acids composed of a cryIB core N-terminal toxin portion of

CC 845 amino acids and a cryIA(b) C-terminal protoxin portion of 382

CC amino acids. The synthetic gene is based on the cryIB and cryIA(b)

CC genes of Bacillus thuringiensis, with codon usage altered to

CC increase expression in plants, particularly maize. The synthetic

CC gene is contained within plasmid pCIB5520 (NRRL B-21972) under

CC control of the PEPC promoter. When the chimeric insecticidal

CC protein gene is expressed in transgenic maize from both PEPC and

CC pith promoters, insecticidal activity is observed against European

CC corn borer (Ostrinia nubilalis). Recombinant microbial strains

CC transformed with the hyFLIB gene can be used in endomocidal

CC formulations for the biological control of Lepidopteran pests.

XX

SQ Sequence 8854 BP; 2202 A; 2580 C; 2189 G; 1883 T; 0 other;

Query Match 45.7%; Score 2637.8; DB 20; Length 8854;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2639; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3127 gaagcgaattcactggccgctgcttttacaacgtcgtgactgggaaacccctggcgttacc 3186

Db 3888 gtaccgaattcactggccgctgcttttacaacgtcgtgactgggaaacccctggcgttacc 3947

Qy 3187 caacttaatgccttgacgacacatcccccttccgagctggcgttaataagcgaagggcc 3246

Db 3948 caacttaatgccttgacgacacatcccccttccgagctggcgttaataagcgaagggcc 4007

Qy 3247 cgcaccgatcgcccttcccaacagttgcgagcctgaatggcgaaatggcgctgatcgcg 3306

Db 4008 cgcaccgatcgcccttcccaacagttgcgagcctgaatggcgaaatggcgctgatcgcg 4067

Qy 3307 tattttctcttacgcacatctgtgcggttattcacaccgcataatggtgcactctcagtaca 3366

Db 4068 tattttctcttacgcacatctgtgcggttattcacaccgcataatggtgcactctcagtaca 4127

Qy 3367 atctgctctgatccgcgcatagtttaagccagccccgcacacccgcacaccccgctgacgcg 3426

Db 4128 atctgctctgatccgcgcatagtttaagccagccccgcacacccgcacaccccgctgacgcg 4187

Qy 3427 ccctgacgggcttgtctgctccccggcatcccgcttacagacaagctgtgaccgctctccggg 3486

Db 5321 |||||ctgttttgcaccagaaacgctggtgaaagtataaagatgctgaagatcagttgggtg 5380

QY 3851 cacgagtggttacatcagaactggatctcaacagcggtaagatccttgagagttttcgcc 3910

Db 5381 cacgagtggttacatcagaactggatctcaacagcggtaagatccttgagagttttcgcc 5440

QY 3911 ccgaagaacgttttccaatgatgagcaccttttaaaagtctgctatgtggcgcggtattat 3970

Db 5441 ccgaagaacgttttccaatgatgagcaccttttaaaagtctgctatgtggcgcggtattat 5500

QY 3971 cccgtattgacgcgggcaagagcaaacactcggtcgcgcgcatacacactattctcagaatgact 4030

Db 5501 cccgtattgacgcgggcaagagcaaacactcggtcgcgcgcatacacactattctcagaatgact 5560

QY 4031 tggttgagtactcaccaagtcaacagaaaaagcatcttacggatggcatgacagtaagagaaat 4090

Db 5561 tggttgagtactcaccaagtcaacagaaaaagcatcttacggatggcatgacagtaagagaaat 5620

QY 4091 tatgcagtgctgccataaccatgagtataaacactcgggccaaacttacttctgcacaaacga 4150

Db 5621 tatgcagtgctgccataaccatgagtataaacactcgggccaaacttacttctgcacaaacga 5680

QY 4151 tcggaggaccgaaggagctaaaccgcttttttgcacaacatggggatcatgtaaactcgcc 4210

Db 5681 tcggaggaccgaaggagctaaaccgcttttttgcacaacatggggatcatgtaaactcgcc 5740

QY 4211 ttgatcgttgggaaccggagctgaatgaagccataccaaaacgacgagcgtgcacaccacga 4270

Db 5741 ttgatcgttgggaaccggagctgaatgaagccataccaaaacgacgagcgtgcacaccacga 5800

QY 4271 tgcctgtagcaaatggcaacaacgcttgcgcgaactattaaactggcgaactacttactctag 4330

Db 5801 tgcctgtagcaaatggcaacaacgcttgcgcgaactattaaactggcgaactacttactctag 5860

QY 4331 ctccccggcaacaaattaaatagactggatggagcgggataaaagtgcaggaccactctgc 4390

Db 5861 ctccccggcaacaaattaaatagactggatggagcgggataaaagtgcaggaccactctgc 5920

QY 4391 gctcggcccttcggctggctggtttattgtctgataaaatctggagccggtgagcgtgggt 4450

Db 5921 gctcggcccttcggctggctggtttattgtctgataaaatctggagccggtgagcgtgggt 5980

QY 4451 ctcgcggtatcatctgcagcactggggccagatggtaagccctcccgatcgtagttatct 4510

Db 5981 ctcgcggtatcatctgcagcactggggccagatggtaagccctcccgatcgtagttatct 6040

QY 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgcctgagataggtg 4570

Db 6041 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgcctgagataggtg 6100

QY 4571 cctcactgattaaagcattggtaactgtcagaccaaagtttactcatatatactttagattg 4630

Db 6101 cctcactgattaaagcattggtaactgtcagaccaaagtttactcatatatactttagattg 6160

QY 4631 atttaaaacttcatttttaatttaaaaggatctaggatgaagatccttttttgataatctca 4690

Db 6161 atttaaaacttcatttttaatttaaaaggatctaggatgaagatccttttttgataatctca 6220

QY 4691 tgacaaaaatcccttaacgtgagttttctgttcactgagcgtcagaccccgtagaaaaa 4750

Db 6221 tgacaaaaatcccttaacgtgagttttctgttcactgagcgtcagaccccgtagaaaaa 6280

QY 4751 tcaaaggatcttcttgagatcccttttttctggcgtaactctgctgcttgcaaacaaaa 4810

Db 6281 tcaaaggatcttcttgagatcccttttttctgcgtaactctgctgcttgcaaacaaaa 6340

QY 4811 aaccaccgtaccagcgggtggtttgtttgcgggatcaagagctaccaactttttccga 4870

Db 6341 aaccaccgtaccagcgggtggtttgtttgcgggatcaagagctaccaactttttccga 6400

QY 4871 aggtaactggcttcacagagcgcagataccaaaatactgtccttctagttagccgtagt 4930

Db 6401 aggtaactggcttcacgagcgcagataccaaaataactgtccttctagtgtagccgtagt 6460

QY 4931 taggccaccacttcaagaactctgtagcacccgcctacatacctcgtctcgttaaatcctgt 4990

Db 6461 taggccaccacttcaagaactctgtagcacccgcctacatacctcgtctcgttaaatcctgt 6520

QY 4991 taccagtggtcgtcgcagtggcgataaagtctgtcttaccgggttggaactcaagacgat 5050

Db 6521 taccagtggtcgtcgcagtggcgataaagtctgtcttaccgggttggaactcaagacgat 6580

QY 5051 agttaccggataaaggcgcagcggctcgggctgaacgggggttctcgtgcacacagcccagct 5110

Db 6581 agttaccggataaaggcgcagcggctcgggctgaacgggggttctcgtgcacacagcccagct 6640

QY 5111 tggagcgaacgacctacacccaactgagatacctacacgctgagctatgagaaaagcgcca 5170

Db 6641 tggagcgaacgacctacacccaactgagatacctacacgctgagctatgagaaaagcgcca 6700

QY 5171 cgcttcccgaagggaagaaaggcggacaggtatccggtaaagcggcagggtcggaaacaggag 5230

Db 6701 cgcttcccgaagggaagaaaggcggacaggtatccggtaaagcggcagggtcggaaacaggag 6760

QY 5231 agcgcacgaggggagcttccaggggggaaacgcctggtatatctttatagtcctcgtcgggttcc 5290

Db 6761 agcgcacgaggggagcttccaggggggaaacgcctggtatatctttatagtcctcgtcgggttcc 6820

QY 5291 gccacctctgacttgagcgtcgatttttctgtgatgctcgtcaggggcgagcctatgga 5350

Db 6821 gccacctctgacttgagcgtcgatttttctgtgatgctcgtcaggggcgagcctatgga 6880

QY 5351 aaaaacgacgcaacgcggcctttttacgggttccctggccttttctggccttttgctcacaca 5410

Db 6881 aaaaacgacgcaacgcggcctttttacgggttccctggccttttctggccttttgctcacaca 6940

QY 5411 tgttcttctcgtgcttatccccctgattctctgtggataaccgtattaccgccttttgagtgag 5470

Db 6941 tgttcttctcgtgcttatccccctgattctctgtggataaccgtattaccgccttttgagtgag 7000

QY 5471 ctgataccgctcgcgcgacgcgaacgacccgagccgagcgcagcagtcagtcagcgggaaagcgg 5530

Db 7001 ctgataccgctcgcgcgacgcgaacgacccgagcgcagcagtcagtcagcgggaaagcgg 7060

QY 5531 aagagcggcccaatacgcgaacccgcctctctccccgcgcttcccgagcttaataatgcagct 5590

Db 7061 aagagcggcccaatacgcgaacccgcctctctccccgcgcttcccgagcttaataatgcagct 7120

QY 5591 ggcacgacaggtttccgcagactggaaaagcgggcagtcagtcagcgaacaaatgaatgtgagtt 5650

Db 7121 ggcacgacaggtttccgcagactggaaaagcgggcagtcagtcagcgaacaaatgaatgtgagtt 7180

QY 5651 agctcactcattaggcaccccccaggcttttacactttatgcttccggtcgtatgtgtg 5710

Db 7181 agctcactcattaggcaccccccaggcttttacactttatgcttccggtcgtatgtgtg 7240

QY 5711 gaattgtgagcgggataacaatttcacacagggaaaacagctatgacctgattacgcca 5767

Db 7241 gaattgtgagcgggataacaatttcacacagggaaaacagctatgacctgattacgcca 7297

RESULT 7
AAD04946/c
ID AAD04946 standard; DNA; 8062 BP.

XX AAD04946;
AC AAD04946;
XX
DT 17-JUL-2001 (first entry)
XX
DE Plasmid pRK76 used to test Cre recombinase mediated inversion.

XX
DE Gene trapping construct; conditional mutation; unidirectional inversion;
XX recombinase recognition sequence; RRS; disruption cassette;
KW selection cassette; transgenic organism; pRK76; Cre recombinase; ds.
KW
XX

OS Chimeric - Adenovirus.
OS Chimeric - ECMV virus.
OS Chimeric - Unidentified.
XX
PN WO200129208-A1.
XX
XX 26-APR-2001.
PD
XX
XX 16-OCT-2000; 2000WO-EP10162.
PF
XX 16-OCT-1999; 99EP-0120592.
PR
PR 27-OCT-1999; 99US-0162016.
XX
XX (ARTE-) ARTEMIS PHARM GMBH.
PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX
XX Kuehn R, Von Melchener H, Altschmied J;
PI
XX
XX WPI; 2001-308486/32.
DR
XX
XX New gene trapping construct capable of causing conditional mutations in
PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped -
XX
XX Example 3; Page 70-73; 78pp; English.
PS
XX The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC recombinase which is capable of unidirectional inversion of double
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is PRK76 vector, which is used to test Cre
CC recombinase mediated inversion, which is related to the invention.
XX
SQ Sequence 8062 BP; 1960 A; 2036 C; 2121 G; 1945 T; 0 other;

Query Match 45.7%; Score 2637; DB 22; Length 8062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactggccgctggttttaacaacgctgtagctgggaaacccctggcgttacccaac 3190
|||||
DB 8045 CGAATTCAGTGGCCGCTCGTTTTACAACTGCTGACTGGGAAACCCCTGGCGTTACCCCAAC 7986

QY 3191 ttaatgccttgacgacacatccccctttcgccagctggcgttaatagcgaagagcccgca 3250
|||||
DB 7985 TTAATCGCCTTGACGACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 7926

QY 3251 ccgatacccttcccaacagttgcgagcctgaatggcgaaatggcctgatgcggtatt 3310
|||||
DB 7925 CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCGCCTGATGCGGTATT 7866

QY 3311 ttctccttacgcatctgtgcgggtatttcacaccgcacatatggtgcactctcagtaacaatct 3370
|||||
DB 7865 TTCTCTCTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCT 7806

QY 3371 gctctgatgccgcatagtttaagccagccccgcacacccgcgaaccccgctgacgcgccct 3430
|||||
DB 7805 GCTCTGATGCCGCATAGTTAAGCCAGCCCCGCACACCCGCCAACACCCGCTGACGCCCT 7746

QY 3431 gacgggttctgtctcccgccatccgcttacagacaagctgtgacccgtctccgggagct 3490
|||||
DB 7745 GACGGGCTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 7686

QY 3491 gcatgtgtcagaggtttttcacocgtcatcaaccgaaacgcgcgagacgaaagggcctcgtga 3550
|||||
DB 7685 GCATGTGTCAGAGGTTTTCACCGTCAACCGAAGCGCGAGAGAAAGGGCCTCGTGA 7626

QY 3551 tacgcctattttataggttaatgtcatgataataatgggtttcttagcgtcaggtggca 3610
|||||
DB 7625 TACGCCATATTTTATAGGTTAATGTCAATGATAAATAATGGTTTCTTAGAGTCAGGTGGCA 7566

QY 3611 cttttcgggaaaaatgtgcgcggaacccctatttggttatttttctaataacattcaata 3670
|||||
DB 7565 CTTTTCGGGAAATGTGCGGGAACCCCTATTGTTTATTATTCTAAATACATTCAAATA 7506

QY 3671 tgtatccgctcatgagacaataaacctgataaaatgcttcaataatattgaaaaaggaaga 3730
|||||
DB 7505 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAGGAAGA 7446

QY 3731 gtagagtattcaacatttcctcgtgtgcgccttattcccttttttgcgcgcattttgccttc 3790
|||||
DB 7445 GTATGAGTATTCAACATTTCCGTTGTCGCCCTTATTCCCTTTTTCGGGCATTTTGCTTC 7386

QY 3791 ctgtttttgctcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 3850
|||||
DB 7385 CTGTTTTTGTCAACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTG 7326

QY 3851 cacgagtggttacatcgaactggatctcaacagcggtaagatccttgagagtttctgcc 3910
|||||
DB 7325 CACGAGTGGTTACATCGAATCGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC 7266

QY 3911 ccgaagaacgttttccaatgatgagcacttttaaagtcttgcctatgtggcgcggtattat 3970
|||||
DB 7265 CCGAAGAACGTTTCCAATGATGAGCACACTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 7206

QY 3971 ccggtattgacgcgcgggcaagagcaactgggtgcgcgcacacattctcagaatgact 4030
|||||
DB 7205 CCGGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAAATGACT 7146

QY 4031 tgggttgactaccagtcacagaaaaagcatcttaccggtggcgtgacagtaagagaat 4090
|||||
DB 7145 TGGTTGAGTACTACCAGTCACAGAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT 7086

QY 4091 tatgcagtgtgccataaccatgagtataaacacactgcggccaacttactctgacaaaga 4150
|||||
DB 7085 TATGCAGTGTGCCATAACCATGAGTGATAACACTCGGGCCAACTTACTTCTGACAAACGA 7026

QY 4151 tcggaggaccgaaggagctaacccgcttttttgcaacaacatgggggatcatgtaactcgcc 4210
|||||
DB 7025 TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAACATGGGGGATCATGTAACTCGCC 6966

QY 4211 ttgatcgttgggaaacccggagctgaatgaagccataccaaaacgacgagcgtgacaccacga 4270
|||||
DB 6965 TTGATCGTTGGGAACCCGGAGCTGAATGAAGCCATACCAACGACGAGCGGTGACACCACGA 6906

QY 4271 tgcctgtagcaatggcaacaacgcttgccgaacactattaaactggcggaactacttactctag 4330
|||||
DB 6905 TGCCTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 6846

QY 4331 ctccccggcaacaattaatagactggatggagggcgataaaagttgcaggaccacttctgc 4390
|||||
DB 6845 CTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACCTTCTGC 6786

QY 4391 gctcggcccttccggcgtggctggtttattgctgataaaatctggagccggtgagcgtgggt 4450
|||||
DB 6785 GCTCGGCCCTTCCGGCTGGCTGCTGTTTATTGCTGTATAAATCTGGAGCCGCTGAGCGTGGGT 6726

QY 4451 ctgcggtatcatcagcactggggccagatggtaagccctcccgtatcgtagttatct 4510
|||||
DB 6725 CTCGCGGTATCATTCAGCACACTGGGGCCACAGATGGTAAGCCCTCCCCGTATCGTAGTTATCT 6666

QY 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg 4570
|||||
DB 6665 ACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 6606

Db 8136 CGAATTCAGTGGCGTGGTTTTACAAAGTCGTGACTGGGAAACCCCTGGCGTTACCCAAC 8077
Qy 3191 ttaatgcgcttgacgacatacccccttccgagctggcgtaataagcgaagagcccgca 3250
Db 8076 TTAATGCGCTTGACGACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 8017
Qy 3251 ccgagccgcttcccaacagttgcgagcctgaatggcggaatggcgccctgatcgcggtatt 3310
Db 8016 CCGATGCGCCTTCCCAACAGTTGCGAGCCTGAATGGCGAATGGCGCCTGATGCGGTATT 7957
Qy 3311 ttctcttacgcatctgtgcggtatttcacacccgcataatggtgcactctcagtacaatct 3370
Db 7956 TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGACCTCTCAGTACAATCT 7897
Qy 3371 gctctgatgcgcatagtttaagccagccccgcagacacccgcccaacacccgctgacgcgccct 3430
Db 7896 GCTCTGATGCCGATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCCCCCT 7837
Qy 3431 gacgggcttgctctgctcccgcatccgccttacagacaagcttgaccgtctccgggagct 3490
Db 7836 GACGGGCTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 7777
Qy 3491 gcatgtgtcagagggttttcacccgcatcaccgaaacgcgcgagacgaaaggcctcgtga 3550
Db 7776 GCATGTGTCAGAGGTTTTTCACCGTCATCACCCGAAACGCCGAGACGAAAGGGCCTCGTGA 7717
Qy 3551 tacgcctatttttataaggttaatgtcatgataataatggttcttagacgtcaggtggca 3610
Db 7716 TACGCCATTTTTATAGGTTAATGTGATGATAATAATATGTTTTCTTAGACGTGAGGTGGCA 7657
Qy 3611 cttttcggggaaatgtgcgcggaaacccctatttgtttatttttctaaatacatcacaata 3670
Db 7656 CTTTTCGGGGAATGTGCGGGAACCCCTATTGTGTTATTGTTTCTAAATACATTCAAATA 7597
Qy 3671 tgtatccgctcatgagacaaataaaccttgataaaatgcttcaataataattgaaaaaggaa 3730
Db 7596 TGTATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATAATTGAAAAAGGA 7537
Qy 3731 gtatgagtattcaaacatttcggtgcgccccctattcccttttttgcgcattttgccttc 3790
Db 7536 GTATGAGTATTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTTGGGCAATTTGCGCTT 7477
Qy 3791 ctgtttttgctcaccagaaacgcgtggtgaaagtataagatgctgaagatcagttgggtg 3850
Db 7476 CTGTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTG 7417
Qy 3851 cacgagtggttacatcgaactgcatctcaacagcggtaagatccttgagagtttgcgc 3910
Db 7416 CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC 7357
Qy 3911 ccgaagaacgtttccaatgatggcaccttttaaagtctgctatgtggcgcgtattat 3970
Db 7356 CCGNAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGCGCGGTATTAT 7297
Qy 3971 cccgtattgacgcgggcaagagcaactcggctcgcgcatacactattctcagaatgact 4030
Db 7296 CCCGTATTGACGCCCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACT 7237
Qy 4031 tgggtgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat 4090
Db 7236 TGGTTGAGTACTCACAGATCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA 7177
Qy 4091 tatgcagtgtgccataaaccatgagtgataaactgcggcccaacttacttctgacaacga 4150
Db 7176 TATGCAGTGTGCCATAAACCATGAGTGATAAACACTGCGGCCAATTACTTCTGACAACGA 7117
Qy 4151 tcggagaccgaaggagcctaaccgtttttgcacaacatgggggatcatgtaactcgcc 4210
Db 7116 TCGGAGACCGAAGGAGCTAACCGCTTTTTCACAACATGSSGGATCATGTAACCTCGCC 7057
Qy 4211 ttgatcgtttgggaaccggagctgaatgaagccataccaaacagcagcgtgacaccacga 4270

Db 7056 TTGATCGTTGGGAACCCGGAGCTGAATGAAGCCATATCCAAACGACGAGCGTGACACCACGA 6997
Qy 4271 tgcctgtagcaaatggcaacaaacgttgcgcaaaactattaactggcgaactacttactctag 4330
Db 6996 TGCCTGTAGCAATGGCAACAACGTTGCCAAAATATTAACTGGCGAACTACTTACTCTAG 6937
Qy 4331 cttcccgccgaacaaattaatagactggatggagcggtataaagtgcaggaaccacttctgc 4390
Db 6936 CTTCCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGACCACTTCTGC 6877
Qy 4391 gctcgcccttccggctggctggtttattgtctgataaaatctggagcgggtgagcgtgggt 4450
Db 6876 GCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAAATCTGGAGCCGCTGAGCGTGGGT 6817
Qy 4451 ctgcggttatcatgtcagcactgggcccagatggtaagccctccctatcgtagttatct 4510
Db 6816 CTGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTTCCCGTATCGTAGTTATCT 6757
Qy 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg 4570
Db 6756 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 6697
Qy 4571 cctcactgattaaagcattggtaactgtcagaccaagtttactcatatataacttttagattg 4630
Db 6696 CCTCACTGATTAAGCATTTGGTAACCTGTAGACCAAGTTTACTCATATATACTTTAGATTG 6637
Qy 4631 atttaaaacttcattttttaaatttaaaaggatctaggtgaagatcctttttgataatctca 4690
Db 6636 ATTTAAAACCTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 6577
Qy 4691 tgacaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaa 4750
Db 6576 TGACCAAAATCCCTTAACGTGAGTTTTTCTCCACTGAGCGTCAGACCCCGTAGAAAAGA 6517
Qy 4751 tcaaaaggatctcttgagatcccttttttctgcgctaactgtcgtcgtctgcaacaaaaa 4810
Db 6516 TCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCTAATCTGCTGCTGCAACAACAAA 6457
Qy 4811 aaccacccgtaccagcgggtgtgtttgttgcgggatcaagagctaccactctttttccga 4870
Db 6456 AACCACCCGTACCAGCGGTGTTGTTGTGCGGGATCAAGAGCTACCAACTCTTTTTCGA 6397
Qy 4871 aggtaaactggcttcagcagagcgcagataaccaaaatactgtccttctagtgtagccgtagt 4930
Db 6396 AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTAGTGTAGCCGTAGT 6337
Qy 4931 taggccacacttcaagaactctgttagcaccgcctacatacctcgtcctgtctaactcctgt 4990
Db 6336 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCCTACATACCTCGCTCTGCTAATCCTGT 6277
Qy 4991 taccagtggctgtccagtggcgataaagtgcgtgtcttacccgggttgactcaagacgat 5050
Db 6276 TACCAGTGGCTGTGCCAGTGGCGATAAGTCGTGTCTTTACCGGGTTGGACTCAAGACGAT 6217
Qy 5051 agttaccggataaaggcgagcgggtcgggtgaacggggggttcgtgcacacagcccagct 5110
Db 6216 AGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGTTCGTCACACAGCCAGCT 6157
Qy 5111 tggagcgaacgacctacacccgaactgagataacctacagcgtgagctatgagaaagcgcca 5170
Db 6156 TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGANAAGCGCCA 6097
Qy 5171 cgcttcccgaaaggagaaaggcggaacaggtatcccgtaagcggcgaggtcggaaacaggag 5230
Db 6096 CGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGTCGGGAACAGGAG 6037
Qy 5231 agcgcaggggagcttccagggggaaacgcctggtatctttatagctcgtcgggttc 5290
Db 6036 AGCGCACGAGGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCTGTGCGGTTTC 5977
Qy 5291 gccacctctgacttgagcgtcgatttttctgatgctcgtcagggggggagcctatgga 5350
Db 5976 GCCACCTCTGACTTGAGCGTCGATTTTTTGTGTGATGCTCGTCAGGGGGGGGAGCCTATGGA 5917

Qy 5590 tggcacgacaggtttcccgactggaagcggggcagtgagcgcaacgcaattaatgtgagt 5649
|||||
Db 9158 tggcacgacaggtttcccgactggaagcggggcagtgagcgcaacgcaattaatgtgagt 9217
|||||
Qy 5650 tagctcactcattaggcacccccaggctttacacatttatgcttccggctcgtatgtgtgt 5709
|||||
Db 9218 tagctcactcattaggcacccccaggctttacacatttatgcttccggctcgtatgtgtgt 9277
|||||
Qy 5710 ggaattgtgagcgggataacaatttcacacagggaacagctatgacctgattacgcca 5767
|||||
Db 9278 ggaattgtgagcgggataacaatttcacacagggaacagctatgacctgattacgcca 9335
|||||
RESULT 11
AAV63742
ID AAV63742 standard; DNA; 6898 BP.
XX
AC AAV63742;
XX
DT 12-APR-1999 (first entry)
XX
DE Plasmid UGP232-4 containing ubi promoter and per5 3'UTR.
XX
KW Peroxidase; per5 gene; maize; corn; transgenic plant; promoter;
KW root; vector; plasmid UGP232-4; ubiquitin; ds; circular; cyclic.
XX
OS Chimeric - Escherichia coli.
OS Chimeric - Zea mays.
OS Chimeric - maize streak virus.
OS Chimeric - Agrobacterium tumefaciens.
OS Chimeric - synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..14
FT /*tag= a
FT /note= "pUC19 polylinker"
FT promoter 15..993
FT /*tag= b
FT /note= "maize ubiquitin promoter"
FT intron 994..2007
FT /*tag= c
FT /note= "ubiquitin intron"
FT misc_feature 2008..2026
FT /*tag= d
FT /note= "synthetic linker"
FT CDS 2027..3835
FT /*tag= e
FT /product= beta-glucuronidase
FT misc_feature 3836..3890
FT /*tag= n
FT /note= "sequence from pKA882"
FT 3'UTR 3897..4262
FT /*tag= o
FT /note= "per5 3'UTR nt 6066-6431"
FT misc_feature 4269..6898
FT /*tag= p
FT /note= "pUC19 sequence"
XX
PN WO9856921-A1.
XX
PD 17-DEC-1998.
XX
PF 10-JUN-1998; 98WO-US11921.
XX
PR 12-JUN-1997; 97US-0049752.
XX
PA (DOWC) DOW AGROSCIENCES LLC.
XX
PI Ainley M, Armstrong K, Belmar S, Folkerts O, Hopkins N;
PI Menke MA, Paredddy D, Petolino JF, Smith K, Woosley A;
XX
DR WPI; 1999-080904/07.
XX

PT New isolated regulatory sequences for transgenic plants - which are
PT derived from the maize root preferential cationic peroxidase protein
PT (per5) gene
XX
PS Example 21; Page 126-129; 150pp; English.
XX
CC This is the nucleotide sequence of UGP232-4, a plasmid comprising
CC a gene cassette in which the following components are operably
CC joined: the maize ubiquitin (ubi) promoter, the GUS gene and the
CC maize root preferential cationic peroxidase per5 gene (see AAV63717)
CC 3' untranslated region (3'UTR). Testing in maize callus indicated
CC that the per5 3'UTR functioned 19% better than nos when the GUS
CC gene was driven by the ubi promoter. This could result from
CC changes in the efficiency of processing or increased stability of
CC the message. The invention relates to isolated regulatory
CC sequences, especially promoter, intron and 3'UTR sequences, of the
CC maize per5 gene. Claimed recombinant gene cassettes comprising
CC per5 regulatory sequences are used to control expression of
CC recombinant genes in selected tissue, especially the root, of
CC transformed plants, particularly maize.
XX
SQ Sequence 6898 BP; 1705 A; 1617 C; 1680 G; 1896 T; 0 other;

Query Match 45.7%; Score 2632.8; DB 20; Length 6898;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2634; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3132 gaattcactggccgtcgttttacaaacgtgctgactgggaaaaacctggcgttacccaact 3191
|||||
Db 4263 gaattcactggccgtcgttttacaaacgtgctgactgggaaaaacctggcgttacccaact 4322
|||||
Qy 3192 taatgccttcgacgacacatccccctttccagctggcgtaatagcgaagagggccgcac 3251
|||||
Db 4323 taatgccttcgacgacacatccccctttccagctggcgtaatagcgaagagggccgcac 4382
|||||
Qy 3252 cgatgccttcaccaacagttgcgcagcctgaatggcgaatggcgctgatcggtattt 3311
|||||
Db 4383 cgatgccttcaccaacagttgcgcagcctgaatggcgaatggcgctgatcggtattt 4442
|||||
Qy 3312 tctccttacgcatctgtgcggtattttcacaccgcgcatatggtgcactctcagtacaatctg 3371
|||||
Db 4443 tctccttacgcatctgtgcggtattttcacaccgcgcatatggtgcactctcagtacaatctg 4502
|||||
Qy 3372 ctctgatgcgcgcatagttaaagccagccccgcgacacccgccaaacccgctgacgcgccctg 3431
|||||
Db 4503 ctctgatgcgcgcatagttaaagccagccccgcgacacccgccaaacccgctgacgcgccctg 4562
|||||
Qy 3432 acgggcttgtctgctccggcatcccgctacagacaagctgtgacgctcctcgggagctg 3491
|||||
Db 4563 acgggcttgtctgctccggcatcccgctacagacaagctgtgacgctcctcgggagctg 4622
|||||
Qy 3492 catgtgtcagaggttttcaccgtcatcacggaacgcgcgagacgaaggccctcgtgat 3551
|||||
Db 4623 catgtgtcagaggttttcaccgtcatcacggaacgcgcgagacgaaggccctcgtgat 4682
|||||
Qy 3552 acgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcac 3611
|||||
Db 4683 acgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggcac 4742
|||||
Qy 3612 ttttcggggaaaaatgtgcgcggaacccctatttgtttatttttctaatacatattcaaatat 3671
|||||
Db 4743 ttttcggggaaaaatgtgcgcggaacccctatttgtttatttttctaatacatattcaaatat 4802
|||||
Qy 3672 gtatccgctcatgagacaataaacctctgataaatgcttcaataatatgaaaaaggaagag 3731
|||||
Db 4803 gtatccgctcatgagacaataaacctctgataaatgcttcaataatatgaaaaaggaagag 4862
|||||
Qy 3732 tatgagtattcaacatttcctggtgcgcctatttcccttttttgcggcattttgccttcc 3791
|||||
Db 4863 tatgagtattcaacatttcctggtgcgcctatttcccttttttgcggcattttgccttcc 4922
|||||
Qy 3792 tgtttttgtcaccaccagaaacgctggtgaaagtataaagatgctgaagatcagttgggtgc 3851
|||||

OS Bacillus sp. SK-1.
XX Synthetic.
FH Key Location/Qualifiers
FT promoter 1..223
FT /*tag= a
FT /note= "specifically claimed promoter region in
FT Claim 1 (see ABA04127)"
XX

PN WO200183787-A1.
XX
XX
PD 08-NOV-2001.
XX
PF 26-APR-2001; 2001WO-JP03607.
XX
PR 27-APR-2000; 2000JP-0128528.
XX
XX
PA (BIOL-) BIOLEADERS CORP.
PA (TAKI) TAKARA SHUZO CO LTD.
XX

PI Sung M, Lee S, Hong S, Seo H;
XX WPI; 2002-066535/09.
DR
XX
XX Promoter sequence effective in Escherichia coli and Bacillus for
PT economic large-scale fermentative production of proteins -
XX
XX Claim 7; Page 55-57; 65pp; Japanese.

XX The present invention describes a DNA sequence (I) comprising promoter
CC activity in Escherichia coli or Bacillus cells, where the promoter is
CC derived from the D-AAT (D-amino acid aminotransferase) gene of
CC Bacillus SK-1. (I) can be used for the high level expression of a
CC foreign gene in a bacterial host for economic and efficient large-scale
CC production of proteins such as enzymes, cytokines and antibodies. The
CC present sequence represents the plasmid pHCE197(II) polynucleotide
CC sequence comprising a specifically claimed promoter region from the
CC present invention.

XX Sequence 3753 BP; 953 A; 912 C; 929 G; 959 T; 0 Other;

Query Match 45.6%; Score 2631.8; DB 24; Length 3753;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3133 aattcaactggccgtcgttttacacgtcgtgactgggaaaaacctggcgttacccaaactt 3192
Dd |||||||
QY 3193 aatgccttgacgacacatcccccttccgacgtggcgtaatagcgaagggccgcacc 3252
Dd |||||||
Dd 1170 aatgccttgacgacacatcccccttccgacgtggcgtaatagcgaagggccgcacc 1229
QY 3253 gatcgcccttcccaacagtgcgcagcctgaatggcgaaatggcgccctgatgcggtattt 3312
Dd |||||||
Dd 1230 gatcgcccttcccaacagtgcgcagcctgaatggcgaaatggcgccctgatgcggtattt 1289
QY 3313 ctccctacgcatctgtgcggtatttcacacccgcataatggtgcaactctcagtacaatctgc 3372
Dd |||||||
Dd 1290 ctccctacgcatctgtgcggtatttcacacccgcataatggtgcaactctcagtacaatctgc 1349
QY 3373 tctgatgcgcgcatagttaagccagccccgcacccccgaacccccgcgtgacgcgccctga 3432
Dd |||||||
Dd 1350 tctgatgcgcgcatagttaagccagccccgcacccccgaacccccgcgtgacgcgccctga 1409
QY 3433 cgggcttgtctgctccccgcacatccgcttacagacaagctgtgacctctccggagctgc 3492
Dd |||||||
Dd 1410 cgggcttgtctgctccccgcacatccgcttacagacaagctgtgacctctccggagctgc 1469
QY 3493 atgtgtcagagggttttcacccgtcatcccgaaacgcgcgcgagagaaaggcctcgtgata 3552
Dd |||||||
Dd 1470 atgtgtcagagggttttcacccgtcatcccgaaacgcgcgcgagagaaaggcctcgtgata 1529

QY 3553 cgctatttttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcact 3612
Dd |||||||
Dd 1530 cgctatttttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcact 1589
QY 3613 tttcggggaaaatgtgcgcggaaacctatttgtttatttttctaaatacatattcaaatatg 3672
Dd |||||||
Dd 1590 tttcggggaaaatgtgcgcggaaacctatttgtttatttttctaaatacatattcaaatatg 1649
QY 3673 tatccgctcatgagacaaataaaccttgataaaatgcttcaataataattgaaaaagggaagagt 3732
Dd |||||||
Dd 1650 tatccgctcatgagacaaataaaccttgataaaatgcttcaataataattgaaaaagggaagagt 1709
QY 3733 atgagtattcaacatttccgctgctgccttattcccttttttgcggcattttgccttcc 3792
Dd |||||||
Dd 1710 atgagtattcaacatttccgctgctgccttattcccttttttgcggcattttgccttcc 1769
QY 3793 gttttgctcaccagaaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgca 3852
Dd |||||||
Dd 1770 gttttgctcaccagaaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgca 1829
QY 3853 cgagtgggttacatcgaaactggatctcaacagcggtgaagatcccttgagagattttcgcccc 3912
Dd |||||||
Dd 1830 cgagtgggttacatcgaaactggatctcaacagcggtgaagatcccttgagagattttcgcccc 1889
QY 3913 gaagaacgttttccaatgatgagcaacttttaaagtctgctatgtggcgcggtattatcc 3972
Dd |||||||
Dd 1890 gaagaacgttttccaatgatgagcaacttttaaagtctgctatgtggcgcggtattatcc 1949
QY 3973 cgtattgacgcggggcaagagcaactcggctgcgcgcacatacatattctcagaatgacttg 4032
Dd |||||||
Dd 1950 cgtattgacgcggggcaagagcaactcggctgcgcgcacatacatattctcagaatgacttg 2009
QY 4033 gttgagtactcaccagtcacagaaaagcatcttacggtggcagatggcagacagtaagagaatta 4092
Dd |||||||
Dd 2010 gttgagtactcaccagtcacagaaaagcatcttacggtggcagatggcagacagtaagagaatta 2069
QY 4093 tgcagtgtgccataaaccatgagtataaacactgcggccaaacttacttctgacaacgatac 4152
Dd |||||||
Dd 2070 tgcagtgtgccataaaccatgagtataaacactgcggccaaacttacttctgacaacgatac 2129
QY 4153 ggaggacggaaggagctaaccgcttttttgcaacaacatgggggatcatgtaactcgccctt 4212
Dd |||||||
Dd 2130 ggaggacggaaggagctaaccgcttttttgcaacaacatgggggatcatgtaactcgccctt 2189
QY 4213 gatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacccacgatag 4272
Dd |||||||
Dd 2190 gatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacccacgatag 2249
QY 4273 cctgtagcaaatggcaacaacgcttgcgcaaacattataactggggaactacttactctagct 4332
Dd |||||||
Dd 2250 cctgtagcaaatggcaacaacgcttgcgcaaacattataactggggaactacttactctagct 2309
QY 4333 tcccggcaacaataatagactggatggaggcggataaaagtgcaggaccacttctgcgc 4392
Dd |||||||
Dd 2310 tcccggcaacaataatagactggatggaggcggataaaagtgcaggaccacttctgcgc 2369
QY 4393 tcggcccttccggctggctggtttattgtctgataaaatctggagccggtgagcgtgggtct 4452
Dd |||||||
Dd 2370 tcggcccttccggctggctggtttattgtctgataaaatctggagccggtgagcgtgggtct 2429
QY 4453 cgcggtatcattgcagcactggggccagatggtgaagccctccgctatcgtagttatctac 4512
Dd |||||||
Dd 2430 cgcggtatcattgcagcactggggccagatggtgaagccctccgctatcgtagttatctac 2489
QY 4513 acgacggggagtcaggcaactatggatgaacgaaaaatagacagatcgctgagataggtgcc 4572
Dd |||||||
Dd 2490 acgacggggagtcaggcaactatggatgaacgaaaaatagacagatcgctgagataggtgcc 2549
QY 4573 tcaactgattaagcattggtaactgtccagaccaaagtttactcatatatacttttagattgat 4632
Dd |||||||
Dd 2550 tcaactgattaagcattggtaactgtccagaccaaagtttactcatatatacttttagattgat 2609

QY 4633 ttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttttgataatctcatg 4692
Dbb 2610 ttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttttgataatctcatg 2669
QY 4693 accaaaaatcccttaacgtgagtttttcgttccactgagcgtcagaccccgtagaaaaagatc 4752
Dbb 2670 accaaaaatcccttaacgtgagtttttcgttccactgagcgtcagaccccgtagaaaaagatc 2729
QY 4753 aaaggatcttcttgagatccctttttttctgcgcgtaaatctgctgttcgcaaaacaaaaaa 4812
Dbb 2730 aaaggatcttcttgagatccctttttttctgcgcgtaaatctgctgttcgcaaaacaaaaaa 2789
QY 4813 ccacgcgtaccagcgggtgttggtttgccgggatacaagagctaccaactctttttccgaag 4872
Dbb 2790 ccacgcgtaccagcgggtgttggtttgccgggatacaagagctaccaactctttttccgaag 2849
QY 4873 gtaactggcttcagcagagcgcagataccaaaatactgtccttctagtgtgacgtagtta 4932
Dbb 2850 gtaactggcttcagcagagcgcagataccaaaatactgtccttctagtgtgacgtagtta 2909
QY 4933 ggcaaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaactcctgtta 4992
Dbb 2910 ggcaaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaactcctgtta 2969
QY 4993 ccagtggctgctgccagtggcgataaagtcgtgtcttaccgggttgactcaagacgataag 5052
Dbb 2970 ccagtggctgctgccagtggcgataaagtcgtgtcttaccgggttgactcaagacgataag 3029
QY 5053 ttaccgggataaaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttg 5112
Dbb 3030 ttaccgggataaaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttg 3089
QY 5113 gagcgaacgacctacacccaactgagatacctacagcgtgagctatgagaaagcgcacag 5172
Dbb 3090 gagcgaacgacctacacccaactgagatacctacagcgtgagcattgagaaagcgcacag 3149
QY 5173 ctcccgaaaggagaaaggcggacaggtatccggtaagcggcgaggtcggaaacaggagag 5232
Dbb 3150 ctcccgaaaggagaaaggcggacaggtatccggtaagcggcgaggtcggaaacaggagag 3209
QY 5233 cgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgc 5292
Dbb 3210 cgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgc 3269
QY 5293 cacctctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatggaaa 5352
Dbb 3270 cacctctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatggaaa 3329
QY 5353 aacgcagcaaacgcggcctttttacgggttcctggccttttgccttttgcacatg 5412
Dbb 3330 aacgcagcaaacgcggcctttttacgggttcctggccttttgccttttgcacatg 3389
QY 5413 ttcttctcgttatccctgattctgtggataaacctgattaccgcctttgagtgagct 5472
Dbb 3390 ttcttctcgttatccctgattctgtggataaacctgattaccgcctttgagtgagct 3449
QY 5473 gatacgcgtcgcgcgacgcgaacgcgagcgcgagtcagtcagtcaggaagcggaa 5532
Dbb 3450 gatacgcgtcgcgcgacgcgaacgcgagcgcgagtcagtcagtcaggaagcggaa 3509
QY 5533 gagcgcccaatacgcgaacccgcctctccccgcgcttggccgattcattaatgcagctgg 5592
Dbb 3510 gagcgcccaatacgcgaacccgcctctccccgcgcttggccgattcattaatgcagctgg 3569
QY 5593 cagcacagggtttcccgactggaaagcgggcagtcgagcgaacgcgaattaatgtgagttag 5652
Dbb 3570 cagcacagggtttcccgactggaaagcgggcagtcgagcgaacgcgaattaatgtgagttag 3629
QY 5653 ctactcattaggcacccccaggctttacactttatgcttccggctcgtatgttgtgga 5712
Dbb 3630 ctactcattaggcacccccaggctttacactttatgcttccggctcgtatgttgtgga 3689
QY 5713 attgtgagcgggataacaattttcacacaggaaacagctatgaccatgattacgccca 5767

Db 3690 attgtgagcgggataacaatttcacaggaacagctatgaccatgattacgccca 3744
RESULT 13
ABA04130
ID ABA04130 standard; DNA; 3755 BP.
XX
AC ABA04130;
XX
DT 28-FEB-2002 (first entry)
XX
DE Plasmid pHCE19(II) polynucleotide sequence SEQ ID NO:4.
XX
KW Promoter; gene expression; fermentation; D-AAT; Bacillus SK-1;
KW D-amino acid aminotransferase; economic; large-scale protein production;
KW circular; ds.
XX
OS Bacillus sp. SK-1.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT promoter 1..225
FT /*tag= a
FT /note= "specifically claimed promoter region in
FT Claim 1 (see ABA04128)"
XX
PN WO200183787-A1.
XX
PD 08-NOV-2001.
XX
PF 26-APR-2001; 2001WO-JP03607.
XX
PR 27-APR-2000; 2000JP-0128528.
XX
PA (BIOL-) BIOLEADERS CORP.
PA (TAKI) TAKARA SHUZO CO LTD.
XX
PI Sung M, Lee S, Hong S, Seo H;
XX
DR WPI; 2002-066535/09.
XX
PT Promoter sequence effective in Escherichia coli and Bacillus for
PT economic large-scale fermentative production of proteins -
XX
PS Claim 7; Page 58-60; 65pp; Japanese.
XX
CC The present invention describes a DNA sequence (I) comprising promoter
CC activity in Escherichia coli or Bacillus cells, where the promoter is
CC derived from the D-AAT (D-amino acid aminotransferase) gene of
CC Bacillus SK-1. (I) can be used for the high level expression of a
CC foreign gene in a bacterial host for economic and efficient large-scale
CC production of proteins such as enzymes, cytokines and antibodies. The
CC present sequence represents the plasmid pHCE19(II) polynucleotide
CC sequence comprising a specifically claimed promoter region from the
CC present invention.
XX
SQ Sequence 3755 BP; 951 A; 914 C; 932 G; 958 T; 0 other;

Query Match 45.6%; Score 2631.8; DB 24; Length 3755;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3133 aattcactggccgtcgtttttacaacgtcgtgactgggaaaaacccctggcttaccacactt 3192
Dbb 1112 aattcactggccgtcgtttttacaacgtcgtgactgggaaaaacccctggcttaccacactt 1171
Qy 3193 aatgccttcagcacatcccccttcccgactggcgtaatagcgaagagccccgcacc 3252
Dbb 1172 aatgccttcagcacatcccccttcccgactggcgtaatagcgaagagccccgcacc 1231
Qy 3253 gatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgctgatcggtatttt 3312

Db 1232 gatcgcccttcccaacagttgcgagcctgaatggcgaaatggcgccctgatcggtatttt 1291
QY 3313 ctcccttacgcatactgtgcggtatttcacacccgcataatggtgcactctcagtacaaatctgc 3372
Db 1292 ctcccttacgcatactgtgcggtatttcacacccgcataatggtgcactctcagtacaaatctgc 1351
QY 3373 tctgatgccgcatagttaagccagccccgcagccccgcacccccgcacccccgcctga 3432
Db 1352 tctgatgccgcatagttaagccagccccgcagccccgcacccccgcctga 1411
QY 3433 cgggcttgctgctcccgccatccgccttacagacaaagctgtgaccgtctccgggagctgc 3492
Db 1412 cgggcttgctgctcccgccatccgccttacagacaaagctgtgaccgtctccgggagctgc 1471
QY 3493 atgtgtcagagggttttcacccgtcatcacccgaaacgcgcgagacgaaaggccctcgtgata 3552
Db 1472 atgtgtcagagggttttcacccgtcatcacccgaaacgcgcgagacgaaaggccctcgtgata 1531
QY 3553 cgcctatttttatagggttaatgtcatgataataatgggtttctttagacgtcaggtggccact 3612
Db 1532 cgcctatttttatagggttaatgtcatgataataatgggtttctttagacgtcaggtggccact 1591
QY 3613 ttccggggaaatgtgcgcggaaacccctattttgtttatttttctaaaaatacattcaaatatg 3672
Db 1592 ttccggggaaatgtgcgcggaaacccctattttgtttatttttctaaaaatacattcaaatatg 1651
QY 3673 tatccgctcatgagacaataaaccttgataaaatgcttcaataatttgaaaaaaggaagat 3732
Db 1652 tatccgctcatgagacaataaaccttgataaaatgcttcaataatttgaaaaaaggaagat 1711
QY 3733 atgagtatccaacattcccggtgctgcgccttattcccttttttgcggaattttgaccttcc 3792
Db 1712 atgagtatccaacattcccggtgctgcgccttattcccttttttgcggaattttgaccttcc 1771
QY 3793 gtttttgcctacccagaacacgctggtgaaagtataaagatgctgaagatcagttgggtgca 3852
Db 1772 gtttttgcctacccagaacacgctggtgaaagtataaagatgctgaagatcagttgggtgca 1831
QY 3853 cgagtggttacatcgaaactggatctcaacagcggtaagatccttgagagttttcgcccc 3912
Db 1832 cgagtggttacatcgaaactggatctcaacagcggtaagatccttgagagttttcgcccc 1891
QY 3913 gaagaacgttttccaatgatgagcaacttttaagttctgctatgtggcgcggtattatcc 3972
Db 1892 gaagaacgttttccaatgatgagcaacttttaagttctgctatgtggcgcggtattatcc 1951
QY 3973 cgtattgacgcgggcaagagcaactcggtcgcgcgcatacacactattctcagaatgacttg 4032
Db 1952 cgtattgacgcgggcaagagcaactcggtcgcgcgcatacacactattctcagaatgacttg 2011
QY 4033 gttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaatta 4092
Db 2012 gttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaatta 2071
QY 4093 tgcagtgctgccataaccatgagtgataaacactgcggccaacttacttctgacaaacgatac 4152
Db 2072 tgcagtgctgccataaccatgagtgataaacactgcggccaacttacttctgacaaacgatac 2131
QY 4153 ggaggaccgaagagctaaaccgcttttttgcaacaatgggggatacatgtataactcgcctt 4212
Db 2132 ggaggaccgaagagctaaaccgcttttttgcaacaatgggggatacatgtataactcgcctt 2191
QY 4213 gatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacccagatg 4272
Db 2192 gatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacccagatg 2251
QY 4273 cctgtagcaatggcaacaacgcttgcgcaactattaaactggcgaactacttactctagct 4332
Db 2252 cctgtagcaatggcaacaacgcttgcgcaactattaaactggcgaactacttactctagct 2311
QY 4333 tcccggcaacaataatagactggatggggcggtataaagttgcagagaccacttctgcgc 4392

Db 2312 tcccggcaacaataatagactggatggagggcggtataaagttgcaggaccacttctgcgc 2371
QY 4393 tcggcccttccggctggctggtttatttctgataaaatcttgagccggtgagcgtgggtct 4452
Db 2372 tcggcccttccggctggctggtttatttctgataaaatcttgagccggtgagcgtgggtct 2431
QY 4453 cgcggtatcatctgacgacactggggccagatggtaaagccctcccgatatcgtagttatctac 4512
Db 2432 cgcggtatcatctgacgacactggggccagatggtaaagccctcccgatatcgtagttatctac 2491
QY 4513 acgacggggagtcaggcaactatggatgaacgaataatagacagatcgctgagataggtgcc 4572
Db 2492 acgacggggagtcaggcaactatggatgaacgaataatagacagatcgctgagataggtgcc 2551
QY 4573 tcaactgattaaagcattggtaactgtcagacccaagtttactcatatatactttagattgat 4632
Db 2552 tcaactgattaaagcattggtaactgtcagacccaagtttactcatatatactttagattgat 2611
QY 4633 ttaaaaacttcatttttaatttaaaaggatctaggtgaagatcccttttttgataaatctcatg 4692
Db 2612 ttaaaaacttcatttttaatttaaaaggatctaggtgaagatcccttttttgataaatctcatg 2671
QY 4693 accaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaagatc 4752
Db 2672 accaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaagatc 2731
QY 4753 aaaggatcttcttgagatcccttttttctgcgcgttaactctgctgcttgcaaaacaaaaa 4812
Db 2732 aaaggatcttcttgagatcccttttttctgcgcgttaactctgctgcttgcaaaacaaaaa 2791
QY 4813 ccaccgctaccagcggtggtttgtttgcggatcaagagctaccaactctttttccgaag 4872
Db 2792 ccaccgctaccagcggtggtttgtttgcggatcaagagctaccaactctttttccgaag 2851
QY 4873 gtaactggcttcagcagagcgcagataaccaatactgtccttcttagttagccgtagtta 4932
Db 2852 gtaactggcttcagcagagcgcagataaccaatactgtccttcttagttagccgtagtta 2911
QY 4933 ggcacacacttcaagaactctgtagcacccgctacatacctcgtctgtctaactcctgtta 4992
Db 2912 ggcacacacttcaagaactctgtagcacccgctacatacctcgtctgtctaactcctgtta 2971
QY 4993 ccagtggtcgtgccagtggcgataaagtcgtgtcttaccgggttggactcaagacgatatag 5052
Db 2972 ccagtggtcgtgccagtggcgataaagtcgtgtcttaccgggttggactcaagacgatatag 3031
QY 5053 ttaccggataagggcgcagcgtgcggctgaacgggggttctcgtgcacacagcccagcttg 5112
Db 3032 ttaccggataagggcgcagcgtgcggctgaacgggggttctcgtgcacacagcccagcttg 3091
QY 5113 gagcgaacgacctacacgaactgagataacctacagctgagctatgagaaaagcgccacg 5172
Db 3092 gagcgaacgacctacacgaactgagataacctacagctgagctatgagaaaagcgccacg 3151
QY 5173 cttcccgaaaggagaaaggcgacaggtatccggtaagcggcagggtcgggaacaggagag 5232
Db 3152 cttcccgaaaggagaaaggcgacaggtatccggtaagcggcagggtcgggaacaggagag 3211
QY 5233 cgcacgagggagcttccagggggaaaacgcctgggtatctttatagtcctcgtcgggttctgc 5292
Db 3212 cgcacgagggagcttccagggggaaaacgcctgggtatctttatagtcctcgtcgggttctgc 3271
QY 5293 cacctctgactgagcgtcgatttttcttacgggttccctggccttttctggccttttctgcacatg 5352
Db 3272 cacctctgactgagcgtcgatttttcttacgggttccctggccttttctggccttttctgcacatg 3331
QY 5353 aacgccagcaacgcggccttttttacgggttccctggccttttctggccttttctgcacatg 5412
Db 3332 aacgccagcaacgcggccttttttacgggttccctggccttttctggccttttctgcacatg 3391
QY 5413 ttctttcctgcgttatccctgatctctgtggataaacgctattaccgcctttgagtgagct 5472
Db 3392 ttctttcctgcgttatccctgatctctgtggataaacgctattaccgcctttgagtgagct 3451

Db 3665 ATGAGTATTCAACAATTTCCGTGTCGCCCTTATTCCCTTTTTCGGCATTTGCGCTTCCT 3606
QY 3793 gtttttgcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgca 3852
Db 3605 GTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 3546
QY 3853 cgagtgggttacatcgaactggatctcaacagcggtgaagatccttgagagtttgcgccc 3912
Db 3545 CGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCC 3486
QY 3913 gaagaacgttttccaatgatgagcactttttaagttctgctatgtggcggtattatcc 3972
Db 3485 GAAGAACGTTTTCNAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGCGGTATTATCC 3426
QY 3973 cgtattgacgcgggcaagagcaactcggtcgcgcgcatacacatttctcagaatgacttg 4032
Db 3425 CGTATTGACCGCGGGAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTG 3366
QY 4033 gttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaatta 4092
Db 3365 GTTGAGTACTCACAGTCAACAGAAAAGCATCTTACGGATGGCATGACATGAAGAGAAATTA 3306
QY 4093 tgcagtgcgccataaaccatgagtataaacactcgcggccaacttacttctgacaaacgatc 4152
Db 3305 TGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAACGATC 3246
QY 4153 ggaggaccgaagggactaaaccgcttttttgcacaacatgggggatcatgtaactcgcctt 4212
Db 3245 GGAGGACCGAAGGAGCTAACCGCTTTTTCGACAACATGGGGATCATGTAACTCGCCTT 3186
QY 4213 gatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacccacgatg 4272
Db 3185 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCAGATG 3126
QY 4273 cctgtagcaaatgaacaaacgcttgcgcaaacattataactggcgaaacttactctagct 4332
Db 3125 CCTGTAGCAATGCCAACACGTTTGGCAACTATTAACTGGCGAACTACTTACTCTAGCT 3066
QY 4333 tcccgccattccgctggctggtttattctgtgataaaatctggagccggtgagcgtgggtct 4452
Db 3065 TCCCGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCCTTCTGCGC 3006
QY 4393 tcggcccttccgctggctggtttattctgtgataaaatctggagccggtgagcgtgggtct 4452
Db 3005 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCCGTTGAGCGTGGGTCT 2946
QY 4453 cgcggtatcattgcagcactggggccagatggtaagccctcccgatcgtagttatctac 4512
Db 2945 CGCGGTATCATTGACGACACTGGGGCCAGATGTAAGCCCTCCCGTATCTAGTTATCTAC 2886
QY 4513 acgacgggagtcaggcaactatggatgaacgaataagacagatcgcctgagataggtgcc 4572
Db 2885 ACGACGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTGCC 2826
QY 4573 tcactgattaaagcattggttaactgtcagaccaaagtttactcataataactttagattgat 4632
Db 2825 TCACTGATTAAAGCATTGGTAAGTGTACAGACCAAGTTTACTCATATATACTTTAGATTGAT 2766
QY 4633 taaaaacttcatttttaatttaaaaggatctaggtgaagatccctttttgataaatctcatg 4692
Db 2765 TTAANAACCTCATTTTAATTTAANAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCATG 2706
QY 4693 accaaaaatcccttaacgtgagtttctgttccactgagcgtcagaccccgtagaaaaagatc 4752
Db 2705 ACCANAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATC 2646
QY 4753 aaaggatcttcttgagatcccttttttctgcgcgtaaatctgctgctgcgaacaaaaaa 4812
Db 2645 AAAGGATCTTCTTGAGATCCTTTTTCGCGGTAATCTGCTGCTGTCGAACAAAAA 2586
QY 4813 ccaccgctaccagcggtggtttgttgcgggatcaagagctaccaactctttttccggaag 4872

Db 2585 CCACCGCTACCAGCGGTGGTTTGTTCGCCGGATCAAGAGCTACCAACTCTTTTCCGAAG 2526
QY 4873 gtaactggttcagcagagcgcagatacaccataactgtccttctagtgtagccgtagtta 4932
Db 2525 GTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTTAGTGAGCCGTAGTTA 2466
QY 4933 ggccaccacttcaagaaactctgtagcacccgcctacaatacctcgtctcgtctaactcctgtta 4992
Db 2465 GGCCACCACCTCAAGAACTCTGTAGCACCGCCTACATACTCGCTCTGCTAATCTCTGTTA 2406
QY 4993 ccagtggctgctgccagtggcgataaagtcgtgtcttaccgggttggtgactcaagacgatag 5052
Db 2405 CCAGTGGCTGCTGCCAGTGGCGATAAAGTCGTCTTACCGGTTGGACTCAAGACGATAG 2346
QY 5053 ttaccgggataaggcgagcggttcgggctgaacggggggttcgtgcacacacagcccagcttg 5112
Db 2345 TTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGSGGGGTTTCGTGCACACACGCCAGCTTG 2286
QY 5113 gagcgaacgacctacacccgaactgagataacctacacgctgagctatgagaaaagcgcacg 5172
Db 2285 GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCGCCACG 2226
QY 5173 cttcccggaaggagaaaagcggacaggtatcccggtaaagcggcaggttcggaacagagagag 5232
Db 2225 CTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAGGTTCGGAACAGGAGAG 2166
QY 5233 cgcacgaggagcttcacagggggaacgcctgtatcttatagtcctgtcggggttctgc 5292
Db 2165 CGCACGAGGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCTCTGGGTTTCGC 2106
QY 5293 cacctctgactgagcgtcgatctttgtgatgctcgtcagggggcgagcctatggaaa 5352
Db 2105 CACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATCGAAA 2046
QY 5353 aacgccagcaacggcgcccttttacgggttccttgcccttttgcggccttttgcctcacatg 5412
Db 2045 AACGCCAGCAACGCGCCTTTTACGGTTCTTGCCCTTTTGTCTGGCCTTTGCTCACATG 1986
QY 5413 ttctttcctgcgttatccctgattctctgtggataaccgtattaccgccttttgagtgagct 5472
Db 1985 TTCTTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCT 1926
QY 5473 gataccgctcgcgcgagccgaaacgacgagcgagcagcagtcagtcagtcgagggaaagcgaa 5532
Db 1925 GATACCGCTCGCCGACAGCCGAACGACCGAGCGCAGCAGTCAGTGAGCGAGGAACGGAA 1866
QY 5533 gagcgcccaatacgcgaacccgcctctccccgcgcgttgccgattcaattaatgcagctgg 5592
Db 1865 GAGCGCCCAATACGAAACCCGCTCTCCCCGCGGTTGGCCGATTTCATTAATGCAGCTGG 1806
QY 5593 cagcacaggtttcccgactggaaaagcgggcagtcagtcagcgaacgaattaatgtgagttag 5652
Db 1805 CACGACAGGTTTCCGACTGGAAGCGGCGAGTCAGCGCAGCAGTCAGTGAGGTTAG 1746
QY 5653 ctcactcattaggcacccccaggcttttacacitttatgcttccggctcgtatgtgtgga 5712
Db 1745 CTCACCTCATTAGGCACCCGAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGA 1686
QY 5713 attgtgagcggataaacaatttcacacaggaacacagctatgacccatgattacgcca 5767
Db 1685 ATTGTGAGCGGATAACAATTTACACACAGGAACAGCTATGACCATGATTACGCCA 1631

RESULT 15
AAT86610/c
ID AAT86610 standard; DNA; 4283 BP.
XX
AC AAT86610;
XX
DT 26-MAR-1998 (first entry)
XX
DE Epidermal growth factor-encoding plasmid DNA pWRG1630.
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:24:39 ; Search time 6064.26 Seconds
(without alignments)
12835.357 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcatgctgcaggtc.....ctatgaccatgattacgccca 5767

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_gss:*
13:	em_gss_hum:*
14:	em_gss_inv:*
15:	em_gss_pln:*
16:	em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match %	Description
1	1280.2	22.2	2080 11 BC001541
2	1055.8	18.3	1070 9 AJ281552
3	986.4	17.1	1013 10 BM438846
4	950.8	16.5	1067 9 AU081137
5	924	16.0	1004 9 AJ281480
6	919.4	15.9	1089 9 AU081124
7	875.8	15.2	935 10 BG838279
8	841	14.6	841 9 AL042026
9	824.6	14.3	854 10 BM438950
10	782.4	13.6	1163 9 AU081044
11	777.8	13.5	800 9 AJ281449
12	772.4	13.4	954 9 AL044364
13	764.2	13.3	990 10 BI489087
14	750.4	13.0	931 10 BI488594
15	743.6	12.9	832 10 BG923768
16	728.6	12.6	842 10 BI687610
17	722.4	12.5	724 9 AL645114

C 18	720.8	12.5	847	12	AZ687169	AZ687169 ENTMP05TF
C 19	715.8	12.4	792	10	BI687057	BI687057 603312478
C 20	711.8	12.3	715	9	AL661706	AL661706 AL661706
C 21	705.4	12.2	707	9	AL656688	AL656688 AL656688
C 22	704	12.2	818	10	BI180988	BI180988 UNL-P-F2-
C 23	701.8	12.2	705	9	AL635845	AL635845 AL635845
C 24	701.8	12.2	705	9	AL643164	AL643164 AL643164
C 25	701.4	12.2	838	12	AZ679307	AZ679307 ENTLG84TF
C 26	698	12.1	718	9	AL631067	AL631067 AL631067
C 27	697.8	12.1	780	10	BI753192	BI753192 603026066
C 28	691	12.0	1039	9	AU081040	AU081040 AU081040
C 29	688.6	11.9	822	9	AI190022	AI190022 qd35e12.x
C 30	686.4	11.9	689	9	AL646532	AL646532 AL646532
C 31	686	11.9	687	9	AU001481	AU001481 AU001481
C 32	680.6	11.8	700	9	AJ281616	AJ281616 4A3A-P8A1
C 33	679	11.8	680	9	AL646751	AL646751 AL646751
C 34	678	11.8	788	9	AJ281697	AJ281697 4A3A-P9E1
C 35	677.8	11.8	757	12	AQ856202	AQ856202 nbeb0001F
C 36	673.2	11.7	754	10	BG920379	BG920379 602826032
C 37	672.4	11.7	675	9	AL636713	AL636713 AL636713
C 38	671	11.6	693	9	AL598527	AL598527 DKF2p3130
C 39	669	11.6	670	9	AL662130	AL662130 AL662130
C 40	668.4	11.6	670	9	AL635952	AL635952 AL635952
C 41	666.4	11.6	669	9	AL660789	AL660789 AL660789
C 42	665.8	11.5	672	9	AL640799	AL640799 AL640799
C 43	665.6	11.5	703	9	AJ281437	AJ281437 4A3A-P4C3
C 44	664.8	11.5	671	9	AL660035	AL660035 AL660035
C 45	662.4	11.5	675	9	AL645322	AL645322 AL645322

ALIGNMENTS

RESULT 1

BC001541

LOCUS BC001541 2080 bp mRNA linear HTC 31-JAN-2002

DEFINITION Homo sapiens, clone IMAGE:3453362, mRNA.

ACCESSION BC001541

VERSION BC001541.1 GI:14705895

KEYWORDS HTC.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2080)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 4 Row: e Column: 19

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7710156

This clone has the following problem: no polyA-tail.

Location/Qualifiers

1. .2080

/organism="Homo sapiens"

FEATURES

source

/db_xref="taxon:9606" /clone="IMAGE:3453362" /tissue_type="Placenta, choriocarcinoma" /clone_lib="NIH_MGC_10" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" BASE COUNT 319 a 692 c 651 g 418 t ORIGIN									
Query Match 22.2%; Score 1280.2; DB 11; Length 2080; Best Local Similarity 82.4%; Pred. No. 2.3e-261; Matches 1638; Conservative 0; Mismatches 3; Indels 346; Gaps 1;									
QY	834	agccatggctccccgcagtgctgctgcacacgccttccctggcttccccactcttct	893						
Db	94	AGCCATGAGGCCCGCAGTGTCTGTGCACACGCCCTTCCCTGGCTTCCCACTCCTTCT	153						
QY	894	cctcctctctggctcctcctgggtgagagagtggggctgagggccgggaggatgcagagct	953						
Db	154	CCTCCTCTCTGGCTCCTGGGTGAGGAGTGGGGCTGAGGCGCGGAGGATGCAGAGCT	213						
QY	954	gctggtgacggtgcgctgggggcccgggctgcggggcattgcctgaagacccccggggccc	1013						
Db	214	GCTGGTGACGGTGCCTGGGGCGCGCTGCGGGGCATTGCGCTGAAGACCCCGGGGGCCC	273						
QY	1014	tgctctgcttctcctgggcatcccttttgcggagccaccatgggacccctgccttct	1073						
Db	274	TGTCCTGCTTCTCTGGGCATCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTCT	333						
QY	1074	gccacggagcccaagcagccttggtcaggggtggtagacgtacaaccttccagagtgt	1133						
Db	334	GCCACGGAGCCCAAGCAGCCTTGCTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGT	393						
QY	1134	ctgctaccaatatgtggacacccctataccagggttttggaggccacccgagatgtggaacc	1193						
Db	394	CTGCTACCAATATGTGGACACCCCTATACCCAGGTTTGGSSCACCCGAGATGTGGAACCC	453						
QY	1194	caacgctgagctgagcaggagactgcctgtacctcaacgtgtggacaccataccccggcc	1253						
Db	454	CAACCGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCC	513						
QY	1254	tacatccccacccctgtcctgtctgtgatctatgggggtggttcttacagtggggcctc	1313						
Db	514	TACATCCCCACCCCTGTCTCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTC	573						
QY	1314	ctccttggaactgtacgatggccgcttcttggtacagggccgagaggaactgtgtgtgc	1373						
Db	574	CTCCTTGGACGTGTACGATGGCCCGCTTCTTGGTACAGGCCGAGAGGACTGTGTGGTGT	633						
QY	1374	catgaactaccgggtgggagcccttggtctcctggcctgcgggggagccgagagccccc	1433						
Db	634	CATGAATAACGGGTGGGAGCCTTGGCTTCTTGCCCTGCGGGGAGCCGAGAGGCCCC	693						
QY	1434	gggcaatgtgggtctcctggatcagaggtgcgccctgcagtgggtgcaggagaactgtgc	1493						
Db	694	GGGCATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAACTGGC	753						
QY	1494	agccttcgggggtgacccgacatcagtgacgctgtttgggagagcgcgggagccgctc	1553						
Db	754	AGCCTTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGGGAGCCGCTC	813						
QY	1554	ggtgggcatgcacctgtgtcccgccccagccgggacctgttccacaggccgctgtgca	1613						
Db	814	GGTGGCATGCACCTGTCTCCCGCCAGCCGGGGCCTGTTCCACAGGGCCGTGTGCA	873						
QY	1614	gagcgtgcccccaatggacccttgggccacggtgggcatgggagaggccctgcagggc	1673						
Db	874	GAGCGTGGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGAGAGGCCCGTCGAGGGC	933						
QY	1674	cacgcagtggccacactgttgggtgtcctccaggcgccactggtgggaatgacacaga	1733						
Db	934	CACGCAGTGGCCCCACCTTGTGGGGTGTCTCTCCAGCGCGGCTGCTGGTGGGAATGACACAGA	993						

QY	1734	gctggtagcctgccttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgt	1793						
Db	994	GCTGTAGCCTGCCCTTCGGACACGACGACGAGGTCCTGTGTGAACACGAATGGCAGT	1053						
QY	1794	gctgcctcaagaaagcgtcttcgggttctccttcgtgcctggttagatggagacttct	1853						
Db	1054	GCTGCCTCAAGAAAGCGTCTTCCGTTCTCCTTTCGTCCCTGTGGTAGATGGAGACTTCT	1113						
QY	1854	cagtgcacacccacagaggccctcatcaaacgcgggagacattccacggccctg	1902						
Db	1114	CAGTGACACCCACAGAGGCCCTCATCAACGGGGGAGACTTCCACGGCCTGCAGGTAAC	1173						
QY	1903		1902						
Db	1174	TGGCTAGCTGGCGTGAAGCTGGCTTCTTGGGTCCCAACGGTCCCTCCCCTCCCTGC	1233						
QY	1903		1902						
Db	1234	GACCCAGGCATGAGGGCTTCTCCAGGCCCATATCCAGAAGTCCACAGAAGTCCCTCC	1293						
QY	1903		1902						
Db	1294	GGCTCAGATCCCAGGGTGGTCAGCAGGGCAGACAGGAAAGCCACCATGGGTCTATT	1353						
QY	1903		1902						
Db	1354	CTTCTCTGCATCCCTCCCCTGATCTCGTCCCTCTCTCTGTCCATGGTTCGGGGTCT	1413						
QY	1903		1902						
Db	1414	CTGTTCAATCTCTCTGGCTCTTTGTCTGTGCCATCTGTCTGTCTACTTGTCTGTCT	1473						
QY	1903		1927						
Db	1474	CTGTGCGTCCATCCCAACCCCTCTTCCCTTACCCCCAGGTGCTGGTGGGTGAAGG	1533						
QY	1928	atgagggctcgtatttctggtttacgggggccccaggcttcagcaagacaaacgagctc	1987						
Db	1534	ATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCAGGCTTACAGAAAGACAAACAGTCT	1593						
QY	1988	tcacagccggggccgagttcctggtccgtgcgggggtgcgggtcggggtccccaggtgacc	2047						
Db	1594	TCATCAGCCGGCGCGAGTTCCTGSCCGGGGTGCGGGTTCGGGTTCGCCAGGTAAGTG	1653						
QY	2048	tggcagccgaggtgctggttcctgcattacacagactggctgcacccagaccggcac	2107						
Db	1654	TGGCAGCCGAGGCTGTGGTCTCTGATACACAGACTGGCTGCTCCGAGGACCCGGCAC	1713						
QY	2108	gcctgagggagggccctgagcgatgtggtggcgaccacaatgctgtgccccgtggccc	2167						
Db	1714	GCCTGAGGGAGGCCCTGAGCGATGTGGTGGCGGACCAACAATGTCTGTGCCCCGTGG	1773						
QY	2168	agctggctggcgactggctgcccagggtgccccgggtctacgctacgtctttgaaacacc	2227						
Db	1774	AGCTGGCTGGCGACTGGCTGCCAGGGTGCCTGGGTCTACGCTACGTCTTTGAACACC	1833						
QY	2228	gtcttccacgctcctcctggccccctgtggatgggggtgccccacgggtacgagatcgagt	2287						
Db	1834	GTGCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGTGGCCACGGCTACGAGATCGAGT	1893						
QY	2288	tcactttgggatccccctggaccctctcgaaactacacggcagagagagaaaattctcg	2347						
Db	1894	TCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGCGAGAGGAGAAATCTTCG	1953						
QY	2348	cccagcgactgatcgatactgggccaacttggcccgacacaggggatccccaatgagcccc	2407						
Db	1954	CCCAGCGACTGATCGGATCTGAGGCCAACTTGGCCCGCACAGGGGATCCCAATGAGCCCC	2013						
QY	2408	gagaccccaaggccccacaatggcccccgctacacggcggggtcagcagtagcttagtc	2467						
Db	2014	GAGACCCCAAGGCCCCACAAATGGCCCCCGGTACACGGCGGGGGGCTCAGCAGTAGCT	2073						

```
Qy 2468 tggacct 2474
Db 2074 TGGACCT 2080

RESULT 2
AJ281552/c
LOCUS
DEFINITION
4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
ACCESSION
AJ281552
VERSION
AJ281552.1 GI:6929432
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE
1 (bases 1 to 1070)
AUTHORS
Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoerge,W., Soares,M.B.
and Kafatos,F.C.
TITLE
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE
20300950
COMMENT
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES
source
Location/Qualifiers
1..1070
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT
263 a 283 c 255 g 269 t
ORIGIN

Query Match 18.3%; Score 1055.8; DB 9; Length 1070;
Best Local Similarity 99.7%; Pred. No. 9.8e-214;
Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 4293 gttgcgcaactattaaactggcgaaactacttactctagcttccgcgcaacaattaataga 4352
Db 1070 GTTGGCGCAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGA 1011

Qy 4353 ctggatgagcgcgataaaggctgcagaccacttctgcgctcgcccttccggctggctg 4412
Db 1010 CTGGATGGAGCGGATAAAGTTGCAGGACCACCTTCTCGCTCGG-CTTCGGCTGGCTG 952

Qy 4413 gtttattgctgataaaatctggagccggtgagcggtgggtctcgcggtatcattgcagcact 4472
Db 951 GTTTATTGCTGATAAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACT 892

Qy 4473 gggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaac 4532
Db 891 GGGGCCAGATGGTAAGCCCTCCCCTATCGTAGTTATCTACACGACGGGAGTTCAGGCAAC 832

Qy 4533 tatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggta 4592
```

```
Db 831 TATGGATGGACGAAATAGACAGATCGCTGAGATAGGTGCCCTCACTGATTAAAGCATTTGTA 772
Qy 4593 actgtcagaccaagtttactactatatactatttagattgatttaaaacttcattttaatt 4652
Db 771 ACTGTCAGACCAAGTTTACTCATATATACTTTTAGATTGATTTAAAACTTCATTTTAAATT 712
Qy 4653 taaaaggatctaggtgaagatcccttttttgataatctcatgaccacaaatcccttaacgtga 4712
Db 711 TAAAAGGATCTAGGTGAAGATCCCTTTTGTATAATCTCATGACCAAAATCCCTTAACGTGA 652
Qy 4713 gttttcgttccactgagcgtcagaccccgtagaaaaagatcaaaaggatcttcttgagatcc 4772
Db 651 GTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCC 592
Qy 4773 ttttttctgcgcgtaaatctgctgcttgcaaaacaaaaaacccacgcgtaccagcgtggt 4832
Db 591 TTTTTCGTCGCGGTAATCTGCTGCTTGCAAAACAAAAAACCCACCGCTACCAGCGGTGGT 532
Qy 4833 ttgtttccggatcaagagctaccacactcttttccgaaggtaactggcttcagcagagc 4892
Db 531 TTGTTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAACCTGGCTTCAGCAGAGC 472
Qy 4893 gcagataccaaaactgtccttcttagttagcgcgtagttaggccaccacttcaagaactc 4952
Db 471 GCAGATACCAAAATACTGTTCTTCTAGTGTAGCCGCTAGTTAGGCCACCACTTCAAGNACTC 412
Qy 4953 tgtagcacgcctacatacctcgtctgtgtaactcctgttaccagtggtgctgccagtg 5012
Db 411 TGTAGCACCGCTACATACCTCGCTCTGCTGTAATCCTGTTACCAGTGGTGTGCCAGTGG 352
Qy 5013 cgataaagtcgtgtcttaccgggttggaactcaagacgatatgttaccggataaaggcgcagcg 5072
Db 351 CGATAAGTCGTCTTACC GGTTGGACTCAAGACGATAGTTACCGGATAAGGCCGACGG 292
Qy 5073 gtcgggctgaacgggggggttcgtgcacacagcccagcttgaggcgaacgacctacaccga 5132
Db 291 GTCGGGCTGAACGGGGGGTTCGTGTCACACAGCCCAGCTTGGAGCGAAGACCTACACCGA 232
Qy 5133 actgagatacctacagcgtgagctatgagaaagcgcacgcttcccgaaaggagaaagc 5192
Db 231 ACTGAGATACTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCGAAAGGGAGAAAGGC 172
Qy 5193 ggacaggtatccggtaagcgagcgaggtcggaacagagagcgacgagggagcttccagg 5252
Db 171 GGACAGGTATCCGGTAAGCGGCGAGGTGCGAACAGGAGAGCGCACGAGGAGCTTCCAGG 112
Qy 5253 gggaaacgctgggtatctttatagtcctgtcgggtttcgcacacctctgacttgagcgtcg 5312
Db 111 GGGAAACGCGCTGGTATCTTTATAGTCTCTCGGGTTTCGCCACCTCTGACTTGAGCGTCG 52
Qy 5313 atttttgcgtgcgtcagggggggggagcgctatggaaaaaacgccagcaa 5363
Db 51 ATTTTGTGATGCTCGTCAGGGGGGGGAGCCCTATGGAAAAACGCCACGAA 1
```

```
RESULT 3
BM438846
LOCUS
DEFINITION
IPLvr00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA
sequence.
ACCESSION
BM438846
VERSION
BM438846.1 GI:18460568
KEYWORDS
EST.
SOURCE
channel catfish.
ORGANISM
Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
REFERENCE
1 (bases 1 to 1013)
AUTHORS
Feng,J., Kucuktas,H., Kocabas,A., Li,P. and Liu,Z.
TITLE
Transcriptome of channel catfish (Ictalurus punctatus): Initial
analysis of expressed sequence tags from the liver
```


JOURNAL
 COMMENT
 Unpublished (2002)
 Contact: Liu ZJ
 The Fish Molecular Genetics and Biotechnology Laboratory,
 Department of Fisheries and Allied Aquacultures and Program of Cell
 and Molecular Biosciences
 Auburn University
 203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
 Tel: 334 844 4054
 Fax: 334 844 9208
 Email: zliu@acesag.auburn.edu
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1. .1013
 /organism="Ictalurus punctatus"
 /db_xref="taxon:7998"
 /clone_lib="Liver cdna library"
 /note="Organ: Liver; Vector: pSport1; Site_1: NotI;
 Site_2: SalI"
 273 a 228 c 245 g 266 t 1 others
 BASE COUNT
 ORIGIN

Query Match		17.1%; Score 986.4; DB 10; Length 1013;	
Best Local Similarity		99.5%; Pred. No. 5.5e-199;	
Matches 1009; Conservative		1; Mismatches	2; Indels
		2; Gaps	2;
QY	3778	gcattttgacctcctgttttttgcaccagaaaacgctgggtgaaagtataaaagatgctgaa	3837
DB	1	GCATTTTGCCCTCCTGTCTTTTGCTCACCAGAGACGCTGCTGAAAGTTAAAGATGCTGAA	60
QY	3838	gatcagttgggtgcacgagtggtttacatcgaactgcatcgaacggttaacagcggttaagatccit	3897
DB	61	GATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTT	120
QY	3898	gagagttttcgccccgaagaacggttttccaatgatgagcaacttttaagttctgctatgt	3957
DB	121	GAGAGTTTTCGCCCCGGAAGAACGTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATGT	180
QY	3958	ggcgcggtattatcccgattatgacgcggggaagagcaactcggtcgcgcatacactat	4017
DB	181	GGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTAT	240
QY	4018	tctcagaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatg	4077
DB	241	TCTCAGAATGACTTGTTGAGTACTCACCAGTCAACAGAAAGCATCTTACGGATGGCATG	300
QY	4078	acagtaagagaattatgcagtgctgccataaccatgagtgataaacactgcggccaactta	4137
DB	301	ACAGTAAGAGAAATTATGCAGTGTCTGCCATAACCATGAGTGATAACACTGCGGCCAACITTA	360
QY	4138	cttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggat	4197
DB	361	CTTCTGACAACGATCGGAGGACCGAA - GAGCTAACCCGTTTTTTTGACAACATGGGGGAT	419
QY	4198	c-atgtaactgcgcttgatcgtttgggaaccggagctgaatgaagccataccaaacgaaga	4256
DB	420	CAATGTAACCTGCCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGA	479
QY	4257	gcgtgacacacagatgcctgtagcaatggcaacaacgttgcgcaaaactattaactggcga	4316
DB	480	CGGTGACACCACGATGCCTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGA	539
QY	4317	actactactctagcttcgcggaacaattaatagactggatggagcggtataaagtgc	4376
DB	540	ACTACTTACTCTAGCTTCCCGGCAACAATTAAATAGACTGGATGGAGCGGATAAAGTTGC	599
QY	4377	aggaccacttctgcgctcggcccttcggctggctggtttattgctgataaaatctggagc	4436
DB	600	AGGACCACCTTCTGCGCTTCGGCCCTTCGGGCTGGCTGTTTATTGCTGATAAATCTGGAGC	659
QY	4437	cggtgagcgtgggtctcgcgggtatcatctgcagcactggggccagatggttaagccctcccg	4496
DB	660	CGGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACACTGGGGCCAGATGGTAAGCCCTCCCG	719

QY	4497	tatcgtagttatctacacgacgggagtcaggcaactatggatgaacgaaatagacagat	4556
DB	720	TATCGTAGTTATCTACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGAT	779
QY	4557	cgctgagataggtgcctcactgattaaagcattggtaaactgtcagaccaagtttactcata	4616
DB	780	CGCTGAGATAGGTGCCCTCACTGATTAAAGCATTTGGTAACCTGTAGACCAAGTTTACTCATA	839
QY	4617	tatacttttagattgatttaaaacttcattttttaatttaaaaggatctaggatgaagatcct	4676
DB	840	TATACTTTAGATTGATTTAAAACCTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCT	899
QY	4677	ttttgataatctcatgacccaaaatccccttaacgtgagtttctcgttccactgagcgtcaga	4736
DB	900	TTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGA	959
QY	4737	ccccgtagaaaagatcaaaaggatcttcttgagatccttttttctgcgcgtaat	4790
DB	960	CCCCGTAGAAAAGATCAAAGGRTCTCTTGAGATCCTTTTTTTCTGCGCGTAAT	1013

RESULT	4
AU081137/c	
LOCUS	
DEFINITION	AU081137 linear EST 30-AUG-2001 AU081137 Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone K12, mRNA sequence.
ACCESSION	AU081137
VERSION	AU081137.1 GI:6431485
KEYWORDS	EST.
SOURCE	rainbow trout.
ORGANISM	Oncorhynchus mykiss Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE	1 (bases 1 to 1067)
AUTHORS	Kono,T., Sakai,M. and LaPatra,S.E.
TITLE	Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus
JOURNAL	Mar. Biotechnol. 2 (5), 493-498 (2001)
COMMENT	Contact: Masahiro Sakai Faculty of Agriculture Miyazaki University 1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES	Location/Qualifiers 1..1067 /organism="Oncorhynchus mykiss" /db_xref="taxon:8022" /clone="K12" /clone_lib="Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus" /tissue_type="Kidney infected by infectious hematopoietic necrosis virus"
BASE COUNT	235 a 292 c 288 g 250 t 2 others
ORIGIN	

		Query Match	16.5%;	Score 950.8;	DB 9;	Length 1067;
		Best Local Similarity	99.1%;	Pred. No. 2.1e-191;		
Matches	977;	Conservative	0;	Mismatches	7;	Indels 2; Gaps 2
QY	4782	gcgcgtaatctgctgttgcaacaaacccccaccggctaccagcggtagtggtttgtttgcc				4841
Dd	1067	GCGCGTAATCTGCTGC-TGCCAACAAAAAAAACACGGCTACCAGC-GTTGTGTTTGGCC				1010
QY	4842	ggatcaagagtactaccaactctttccgaaggtaaactggcttcagcagagcgcagatacc				4901
Dd	1009	GGATCAAGAGCTACCAAACCTTTTTTCGGAAGGTAACGTGGCTTCAGCAGAGCGCATACC				950
QY	4902	aaatactgtcccttctagtgtagcogtagttaqgccaccacttccaagaacctctgtatqcacc				4961


```
|||||
Db 949 AAATACTGTCCTTCTAGTGTAGCCGTACTTAGCCACCACCTTCAAGAACTCTGTAGCACC 890
QY 4962 gcctacatacctcgctctgctaatacctgttaccagtggtgctgctgcagtggcgataagtc 5021
Db 889 GCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTC 830
QY 5022 gtgtcttaccgggttgactcaagacagatagttaccggataaaggcgagcggtcggtg 5081
Db 829 GTGCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGAGCGGTCTGGGCTG 770
QY 5082 aacggggggttcgtgcacacagcccagcttgagcgaaacgacctacacggaactgagata 5141
Db 769 AACGGGGGTTGCTGCACACAGCCACGCTTGAGCGGAACGACCTACACCGAACTGAGATA 710
QY 5142 cctacagcgtgagctatgagaaagcgccacgcttcccgaaggagaaagcgagagta 5201
Db 709 CCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTA 650
QY 5202 tccggtaagcggcagggtcggaaacagagagcgacgagggagcttccagggggaacgc 5261
Db 649 TCCGGTAAGCGCAGGGTCGGAAACAGGAGCGCACGAGGGAGCTTCCAGGGGGAACGC 590
QY 5262 ctggtatctttatagtcctgctcgggttccgacctctgacctgagcgtcgattttgtg 5321
Db 589 CTGGTATCTTTATAGTCTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTG 530
QY 5322 atgctcgtcagggggcgagcctatgaaaaaacgccagcaacgcgcgcctttttacggtt 5381
Db 529 ATGCTCGTCAGGGGGCGGAGCCTATGGAABACGCCAGCCAGCAACGCGGCTTTTTACGGTT 470
QY 5382 cctggccttttctggccttttgcacatgttcttccctgcgttatccctgattctgt 5441
Db 469 CCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTCGGTTATCCCTGATTCTGT 410
QY 5442 ggataaacctattaccgccttttgagtgcgtgataccgctgcgcgagcggaacgcga 5501
Db 409 GGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCGACCGCAACGCGA 350
QY 5502 gcgcagcagtcagtgagcaggaagcggaagagcgcccaatacgcacaaacgcctctccc 5561
Db 349 CGCGACGGAGTCAGTGAGCGGAGGAAGCGGAAGAGCGCCCCAATACGCCAACCGCCTCTCCC 290
QY 5562 cgcgcgttggccgattcattaatgcagctggcagcacaggtttcccgactggaaagcggg 5621
Db 289 CGCGCGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGG 230
QY 5622 cagtgagcgaacgcaattaatgtgagttagctcactcattaggcaccagcgtttaca 5681
Db 229 CAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAGGCACCCAGGCTTTACA 170
QY 5682 ctttatgcttcgggctcgatgttgttggaattgtgagcggataacaatttcacacagg 5741
Db 169 CTTTATGCTTCGGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGG 110
QY 5742 aaacagctatgaccatgattacgcca 5767
Db 109 AAACAGCTATGACCATGATTACGCCA 84
```

```
RESULT 5
AJ281480/c
LOCUS AJ281480 1004 bp mRNA linear EST 30-JUN-2000
DEFINITION 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cdna clone 4A3A-P4G8, mRNA sequence.
ACCESSION AJ281480
VERSION AJ281480.1 GI:6929360
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
```

```
1 (bases 1 to 1004)
Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansoerge,W., Soares,M.B.
and Kafatos,F.C.
Anopheles gambiae pilot gene discovery project: Identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1..1004
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4G8"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cdna. The 4A3A is a directionally cloned and normalized
cdna library that was constructed from the 4A3A cell line
oligo-T primed cdna according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 252 a 262 c 244 g 244 t 2 others
ORIGIN
Query Match 16.0%; Score 924; DB 9; Length 1004;
Best Local Similarity 99.0%; Pred. No. le-185;
Matches 1001; Conservative 1; Mismatches 2; Indels 7; Gaps 7;
QY 4382 cacttctgcgtcgcggcccttccggctggctggtttattgctgataaaatcggagccggtg 4441
Db 1004 CACTTCTGCGCTCGG-CCTTCGGCTGGCTGGTGTATTGTCGA-AAATCTGGAGCC-GTG 948
QY 4442 agcgtgggtctcgcgggtatcattgcagcactggggccagatggtaagccctcccgatcg 4501
Db 947 AGCGT-GGTCCTCGCGTATCATTTGCAGCACT-GGGCCARATGGTAAGCCCTCCCGTATCG 890
QY 4502 tagttatctacacgacggggagtcaggcaactatggtgaacgaaaatagacagatcgctg 4561
Db 889 TAG-TATCTACACGACGGGAGTTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTG 831
QY 4562 agataggtgcctcactgattaagcattgtaactgtcagaccaagtcttactatataac 4621
Db 830 AGATA-GTGCCTCACTGATTAAAGCATTTGGTAAGTGTGAGACCAAGTTTACTCATATATAC 772
QY 4622 tttagattgatttaaaacttcatTTTTtaatttaaaaggatctaggtgaagatccttttg 4681
Db 771 TTTAGATTGATTTAAAAACTTCATTTTAAATTTAAAAAGGATCTAGGTGAAGATCCTTTTG 712
QY 4682 ataatctcatgacccaaaaatcccttaacgtgagttttcgttccactgagcgtcagacccccg 4741
Db 711 ATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGGTCAGACCCCCG 652
QY 4742 tagaaaagatcaaaaggatcttcttgagatcctttttctgcgcgtaaatctgctgttcg 4801
Db 651 TAGAAAAGATCAAAAGGATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGC 592
QY 4802 aaacaaaaaaaccccgctaccagcgggtggtttgtttgctcgggatcaagagctaccaactc 4861
Db 591 AAACAAAAAAACCCACCGCTACCAGCGGTGGTTGTGTTGCCGGATCAAGAGTACCAACTC 532
QY 4862 ttttccgaaagtaactggcttcagcagagcgagataccacaaatactgtccttctagtgt 4921
Db 531 TTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCGAGATACCAATACTGTTCTTCTTAGTGT 472
```


BG838279
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Glycine clandestina.
Glycine clandestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 935)
Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A., Harris ,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker ,N.A.
Expressed Sequence Tags from Cold-Stressed Glycine clandestina Seedlings
Unpublished (2001)
Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
Location/Qualifiers.
1. 935
/organism="Glycine clandestina"
/cultivar="1035"
/db_xref="taxon:45687"
/clone_lib="Gc01_10e07"
/clone_lib="Gc01_AAFc_ECORC_cold_stressed_Glycine_clandestina"
/tissue_type="Leaves, stem"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI; Site_2: XhoI; Plants incubated at 2 degrees under 12 hours of light/day. Harvested after only 2-3 days of cold treatment. cDNA was prepared with the Uni-Zap cDNA kit from Stratagene. Eco RI adapters were linked followed by digest with Xho I/Eco RI and ligated to pBluescript."
BASE COUNT
243 a 213 c 227 g 238 t 14 others
ORIGIN
*Query Match
Best Local Similarity
Matches
882; Conservative
14; Mismatches
6; Indels
1; Gaps
1;
Qy 3593 cttagacgtcagggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttt 3652
Db 26 CTTACAAATTAGGTGGCACTTTTCGGGAAATGTGCGCGAACCCTATTGTTATTTT 85
Qy 3653 tctaaatacattcaaatatgtatccgctcatgagacaataaacctgataaaatgcttcaat 3712
Db 86 TCTAAATACATTCAAAATATGTATCCGCTCATGAGACAATAACCCGTATAAATGCTTCAAT 145
Qy 3713 aatattgaaaaaggaagagtatgagtattcaacatttcctgtgcgccctattccctttt 3772
Db 146 AATATTGAAAAAGGRAGAGTATGAGTATTCAACATTTCCGTGTGCGCCTTATTCCTTTT 205
Qy 3773 ttgcggcattttgccttcctcctgttttttgctcaccagaaaacgctggtgaaagtaaaagatg 3832
Db 206 TTGCGGCATTTTGCCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATG 265
Qy 3833 ctgaagatcagttgggtgcacgagtggtgttacatcgaaactggatcgaacgagcggtgaaga 3892
Db 266 CTGRAGATCAGTTGGGTGCACGAGTGGGKTACATCGAACTGGATCTCAACAGCGGTAAGA 325
Qy 3893 tccttgagagttttcgccccgaagaacgcttttccaatgatgagcacttttaagtctgc 3952
Db 326 TCCTTGAGAGTTTTCGCCCCGAAGAACGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGC 385

QY 3953 tatgtggcgggtattatcccgtattgagccggggcaagagcaaaactcgggtcgcgcatac 4012
Db 386 TATGTGGCGGGTATTATCCCGTATTGACGCGGGGCAAGAGCAACTCGGTGCGCGCATAC 445
QY 4013 actattctcagaatgacttggttgagtactcaccagtcacagaaaaagcatcttacggatg 4072
Db 446 ACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATG 505
QY 4073 gcatgacagtaagagaattatgcagtgctgccataaaccatgagtataaacactgcgg-cc 4131
Db 506 GCATGACAGTAAGAGAATTATGCAGTCTGCCATAACCATGAGTGATAACACTGCGGCC 565
QY 4132 aacttacttctgacaacgatcgggagggaccgaaggagctaaaccgcttttttgcacaacatg 4191
Db 566 AACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAAACATG 625
QY 4192 ggggatcatgtaactcgcccttgatcgttgggaaccggagctgaatgaagccataccaaac 4251
Db 626 GGGGATCATGTAACTCGCCYTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAA 685
QY 4252 gacgagcgtgacaccacgatgcctgtagcattgcaaatggcaaacgcttgcaaaactattaact 4311
Db 686 GACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACT 745
QY 4312 ggcgaactacttactcttagcttcccgccggcaacaattaatagactggatggagcggtataa 4371
Db 746 GCGGRACTACTTACTTCTAGCTTCCCGGCAACAATAATAGACTGGATGGRGCGGATAAA 805
QY 4372 gttgcaggaccacttctgcgctcggcccttcggctggctggtttatttgcgtgataaatct 4431
Db 806 GWTGCAGGACCACCTTCTGCGCTCGGCCCTTCCGGCTGGSTGGKTTATTCTGTATAWATCT 865
QY 4432 ggagccggtgagcgtgggtctcgcggttatcattgcagcactggggccagatggtaagccc 4491
Db 866 GGRGCCGGKGAGCGTGGGKCTCGCGGTATCATTTGCAGCMCTGGGGGCAGATGGTAAGCCC 925
QY 4492 tcc 4494
Db 926 CCC 928
RESULT 8
AL042026
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
AL042026
DKFZp434E111_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFZp434E111 5', mRNA sequence.
AL042026
AL042026.1 GI:5421372
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 841)
Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
EST (Poustka, et al.)
Unpublished (1999)
Contact: Poustka A.J.
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpimg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp434E111) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. .841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434E111"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: Sali"
BASE COUNT 226 a 192 c 202 g 221 t
ORIGIN

Query Match 14.6%; Score 841; DB 9; Length 841;
Best Local Similarity 100.0%; Pred. No. 4.2e-168;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3639 tatttgttattttctaaatacatattcaattcgaatgtatccgctcatgagacaataaccctg 3698
|||||
Db 1 TATTGTGTTATTTTCTAATAATACATTCAATAATATGTATCCGCTCATGAGACATAACCCGTG 60
|||||

QY 3699 ataaatgcttcaataatattgaaaaaggaagagatgatattcaacatttcggtgtcgc 3758
|||||
Db 61 ATAAATGCTTCAATAATATTGAAAAGGAAGAGATGAGTATTCAACATTCCGTGTGCGC 120
|||||

QY 3759 ccttattcccttttttgcggcattttgccttctctgttttgcacccagaacgctggt 3818
|||||
Db 121 CCTTATTCCTTTTTTGCGGCATTTTGCCCTTCCTGTTTGTCTACCCAGAACGCTGGT 180
|||||

QY 3819 gaaagtaaaagatgctgaagatcagttgggtgcacgagtggttacatcgaaactggatct 3878
|||||
Db 181 GAAAGTAAAGATGCTGAAGATCAGTTGGTGGCAGGAGTGGTTACATCGAACTGGATCT 240
|||||

QY 3879 caacagcggtaagatccttgagagttttgcgccgaagaacggttttccaatgatgagcac 3938
|||||
Db 241 CAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCGCGAAGAACGTTTTCCTCAATGATGAGCAC 300
|||||

QY 3939 ttttaaagttctgctatgtggcgcggtattattcccgctattgacgccgggcaagagcaact 3998
|||||
Db 301 TTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACT 360
|||||

QY 3999 cggtcgcgcgcatacactattctcagaatgacttggttgagttgactcaccagtcacagaaaa 4058
|||||
Db 361 CGGTGCCGCATACACTATTCTCAGAAATGACTGGTTGAGTACTCACCACTCACAGAAA 420
|||||

QY 4059 gcatcttacggatgggcatgacagtaagagaaattatgcagtgctgccataaccatgagtga 4118
|||||
Db 421 GCATCTTACGGATGGCATGACAGTAAGAGAAATATGCAGTGTGCTGCCATAACCATGAGTGA 480
|||||

QY 4119 taacactcggcccaacttactctgacaaacgatcggagaccgaagagagctaacgccttt 4178
|||||
Db 481 TAACACTGCGGCCAACTTACTTCTGACAAACGATCGGAGGACCGAAGGAGCTAACCGCTTT 540
|||||

QY 4179 ttgcaacaacatgggggatcatgtaactcgccttgatcgtttgggaaccggagctgaatga 4238
|||||
Db 541 TTTGCACAACATGGGGGATCATGTAACCTCGCTTGATCGTTGGGAACCGAGCTGAATGA 600
|||||

QY 4239 agccataccaaaacgacgagcgtgacacccagatgcctgtagcaaatggcaacaacgcttgcg 4298
|||||
Db 601 AGCCATACCAACACGACGCGTGACACCCAGATGCCTGTAGCAATGGCAATGGCAACACGTTGCG 660
|||||

QY 4299 caaactattaaactggcggaactacttactctagcttccccgggcaacaattaatagactggat 4358
|||||
Db 661 CAAACTATTAACTGGCGGAACACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGAT 720
|||||

QY 4359 ggagcgggataaagttgcaggaccacttctgcgctcgcccttcggctcggtggtttat 4418
|||||
Db 721 GGAGGCGGATAAAGTTGCAGGACCACCTTCTGCGCTCGGCCCTTCCGGCTGGCTGTTTAT 780
|||||

QY 4419 tgctgataaaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggcc 4478
|||||

|||||
Db 781 TGCTGATAAATCTGGAGCCGGTGAGCGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCC 840
|||||

QY 4479 a 4479
|
Db 841 A 841
|||||

RESULT 9
BM438950
LOCUS
DEFINITION
IPLvr00491 Liver cdna library Ictalurus punctatus cdna 5', mRNA
sequence.
ACCESSION
BM438950
VERSION
BM438950.1 GI:18460672
KEYWORDS
EST.
SOURCE
channel catfish.
ORGANISM
Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
REFERENCE
1 (bases 1 to 854)
AUTHORS
Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Llu, Z.
TITLE
Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
JOURNAL
Unpublished (2002)
COMMENT
Contact: Llu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zllueacesag.auburn.edu
Seq primer: M13 Reverse.
Location/Qualifiers
1. .854
/organism="Ictalurus punctatus"
/db_xref="taxon:7998"
/clone_lib="Liver cdna library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: Sali"
BASE COUNT 226 a 198 c 221 g 209 t
ORIGIN

Query Match 14.3%; Score 824.6; DB 10; Length 854;
Best Local Similarity 99.3%; Pred. No. 1.3e-164;
Matches 849; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 3778 gcattttgccttcctctgttttgcaccaccagaacgctggtgaaaagtaaaagatgctgaa 3837
|||||
Db 1 GCATTTTGCCCTTCCTGTTTGTCTCACCCAGATACGCTGGTGAAAGTAAAGATGCTGAA 60
|||||

QY 3838 gatcagttgggtgcacgagtgagg-ttacatcgaaactggatctcaacagcggttaagatcct 3896
|||||
Db 61 GATCAGTTGGGTGCACGAGTGGGCTTACATCGAACTGGATCTCAACAGCGGTAAGATCCT 120
|||||

QY 3897 tgagagttttgcgccgaagaacggttttccaatgatgagcacttttaagttctgctatg 3956
|||||
Db 121 TGAGAGTTTTCGCCCGGAAGAACGTTTTCCTCAATGATGAGCACTTTTAAAGTTCTGCTATG 180
|||||

QY 3957 tggcgcggtattatcccgctattgacgccgggcaagagcaactcggtcgccgcatacacta 4016
|||||
Db 181 TGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTA 240
|||||

QY 4017 ttctcagaatgacttggttgagtactcaccagtcacagaaaaagcatcttacggatggcat 4076
|||||
Db 241 TTCTCAGAATGACTTGGTTGAGTACTCACCACTCACAGAAAAGCATCTTACGGATGGCAT 300
|||||

QY 4077 gacagtaagagaattatgcagtgctgccataaccatgagtataacactcgggccaactt 4136
|||||

Db 301 GACAGTAAGAGAATTATGCAGTGTGCTGCCATAACCATGAGTGATATAACACTGCGGCCAACTT 360

QY 4137 actctgacaacgacgagcgagaccgaaggagctaacccgctttttgcaacaacatggggga 4196
|||||

Db 361 ACTTCTGACAACGATCGGAGGACCGAA-GAGCTAACCGCTTTTGTGCACAACATGGGGGA 419

QY 4197 tcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacga 4256
|||||

Db 420 TCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGA 479

QY 4257 gcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaaactattaactggcga 4316
|||||

Db 480 GCCTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCCAAACTATTAACTGGCGA 539

QY 4317 actacttactctagcttcccggaacaattaatagactggatggagcggtataaagtgc 4376
|||||

Db 540 ACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGC 599

QY 4377 aggaccacttctgcgctcgcccttcccggttatcattgcagcactggggccagatggttaagccctccg 4496
|||||

Db 600 AGGACCACCTTCTCGGCTCGGCCCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGC 659

QY 4437 cgtgagcgtgggtctcggtcggttatcattgcagcactggggccagatggttaagccctccg 4496
|||||

Db 660 CGGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCAC TGGGCCAGATGGTAAGCCCTCCCG 719

QY 4497 tategtagttatctacacgacggggagtcaggcaactatggatgaacgaatagacagat 4556
|||||

Db 720 TATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAGATAGACAGAT 779

QY 4557 cgctgagatagggtgcctcaactgattaagcattggtaactgtcagaccacgaatttactcata 4616
|||||

Db 780 CGCTGAGATAGGTGCCTCACTGATTAAAGCATTTGGTAACTGTCTGACAGCAAGTTTACTCATA 839

QY 4617 tatacttttagattga 4631
||| ||| ||| ||| |||

Db 840 TATGCTTGAGATTGA 854

RESULT 10

AU081044/c

LOCUS

DEFINITION AU081044 Oncorhynchus mykiss 1163 bp mRNA linear EST 30-AUG-2001
hematopoietic necrosis virus Onco^{rh}ynchus mykiss cdna clone KBl,

ACCESSION AU081044

VERSION AU081044.1 GI:6431392

KEYWORDS EST.

SOURCE rainbow trout.

ORGANISM Onco^{rh}ynchus mykiss

REFERENCE 1 (bases 1 to 1163)

AUTHORS Kono,T., Sakai,M. and LaPatra,S.E.

TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Onco^{rh}ynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus

JOURNAL Mar. Biotechnol. 2 (5), 493-498 (2001)

COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: m.sakai@cc.miyazaki-u.ac.jp.

FEATURES

source

1. .1163
/organism="Onco^{rh}ynchus mykiss"
/db_xref="taxon:8022"
/clone="KBl"
/clone_lib="Onco^{rh}ynchus mykiss Kidney infected by infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic necrosis virus"

BASE COUNT 272 a 307 c 298 g 286 t

ORIGIN

Query Match 13.6%; Score 782.4; DB 9; Length 1163;
Best Local Similarity 99.7%; Pred. No. 1.4e-155;
Matches 794; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 4972 ctcgctctgctaatacctgttaccagtggtgctgccagtgccagtgccgataagtcgtgtcttacc 5031
|||||

Db 1163 CTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACC 1104

QY 5032 gggttggaactcaagacgatagttaccgggataaaggcgcagcgggtcgggctgaacgggggt 5091
|||||

Db 1103 GGGTTGGACTCAAGACGATAGTTACCGGATAAAGCGCAGCGGTGCGGCTGAAC-GGGGGT 1045

QY 5092 tcgtgcacacagcccagcttggagcgaaacgacctacacccgaactgagatacctacagcgt 5151
|||||

Db 1044 TCGTGCACACAGCCCAGCTTGGAGCGAAGACCTACACCGAACTGAGATACCTACAGCGT 985

QY 5152 gagctatgagaaaagcgccacgcttccgaaggagaaaaggcgacaggtatccggtaaagc 5211
|||||

Db 984 GAGCTATGAGAAAAGCGCCACGCTTCCCGAAGGGAGAAAAGCGGACAGGTATCCGGTAAGC 925

QY 5212 ggcagggtcggaaacaggagagcgcacgagggttccagggggaaacgcctggtatctt 5271
|||||

Db 924 GGCAGGTGCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCCTGTTCTT 865

QY 5272 tatagtcctgtcgggttttcgccacctctgacctgagcgtcgatgttttctgtagtcgtca 5331
|||||

Db 864 TATAGTCCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCA 805

QY 5332 gggggcgagacctatggaaaaacgcccagcaacgcgccctttttacgggttccctggcctt 5391
|||||

Db 804 GGGGGCGGAGCCTATGGAAAAACGCCAGAACGCGGCCCTTTTACGGTTCCTGGCCTTT 745

QY 5392 tgttggtcctttgtctcacatgttcttctcgtcgttatccccctgattctgtggataaccgt 5451
|||||

Db 744 TGCTGGCCTTTTGCTCACATGTTCTTCTCTGCGTTATCCCTGATTCTGTGGATAACCGT 685

QY 5452 attaccgcttttgagtgagctgataccgctcgcgcgcgcagccgaaacgagcgcagcgag 5511
|||||

Db 684 ATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCGAGCCGGAACGACGCGCAGCGAG 625

QY 5512 tcagtgagcgagggaagcgggaagagcgcgcccaataacgaaacgcctctccccgcgcgttg 5571
|||||

Db 624 TCAGTGAGCGGAGGAAGCGGAAGAGCGGCCCAATACGAAACCGCCTCTCCCCGCGTGG 565

QY 5572 ccgattcataatgcagctggcagcacaggtttcccgactggaaagcgggcagtgagcgc 5631
|||||

Db 564 CCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCGACTGGAAAGCGGGCAGTGAGCGC 505

QY 5632 aacgcaatgaatgtgagttagctcactcattagggcacccccaggctttacactttatgctt 5691
|||||

Db 504 AACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCGAGGCTTTACACTTTATGCTT 445

QY 5692 ccggctcgtatgtgtgtggaattgtgagcgggataaacaatttcacacaggaaacagctat 5751
|||||

Db 444 CCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTTCACACAGGAAACAGCTAT 385

QY 5752 gaccatgattacgcca 5767
|||||

Db 384 GACCATGATTACGCCA 369

RESULT 11

AJ281449/c

LOCUS

DEFINITION AJ281449 800 bp mRNA linear EST 30-JUN-2000
4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cdna clone 4A3A-P4D5, mRNA sequence.

ACCESSION AJ281449

VERSION AJ281449.1 GI:6929329

KEYWORDS EST.

SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE 1 (bases 1 to 800)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C., Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B. and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.
FEATURES
source Location/Qualifiers
1..800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4D5"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 203 a 206 c 198 g 193 t
ORIGIN
Query Match 13.5%; Score 777.8; DB 9; Length 800;
Best Local Similarity 99.1%; Pred. No. 1.1e-154;
Matches 782; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 4620 actttagattgattaaacttcatttttaattaaaggatctaggtgaagatcctttt 4679
Db 800 ACTTTAGATTGATTAAAACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCTTTT 741
Qy 4680 tgataatctcatgaccacaaatcccttaacgtgagttttcgttccactgagctcagaccc 4739
Db 740 TGATAATCTCATGACCACAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCC 681
Qy 4740 cgtagaaaagatcaaaggatcttcttgagatcccttttttctgagcgttaactgtgctt 4799
Db 680 CGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGGTAACTGCTGCTT 621
Qy 4800 gcaaacacacacacacccgctaccagcgtgtgtttgtttgcccgatcaagagctaccaac 4859
Db 620 GCAAACAAAAAAACCCACCGCTACCGCGTGGTTTGTGTTGCCGGATCAAGAGCTACCAAC 561
Qy 4860 tcttttccgaaggtaactggcttcagcagagcgagataccaaatactgtccttctagt 4919
Db 560 TCTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCTTCTAGT 501
Qy 4920 gtagccgtagttaggccaccacttcaagaactctgtagcaccgctacatacctcgctct 4979
Db 500 GTAGCCGTAGTTAGGCCACCACCTTCAAGAACCTCTGTAGCACCSCCTACATACCTCGCTCT 441
Qy 4980 gctaactcgtttaccagtggtgctgcccagtgggcgataagtcgtgtcttaccgggttga 5039
Db 440 GCTAATCTGTTACCAGTGGCTGCCAGTGGCGGATAAGTCGTGTTCTTACCGGGTTGGA 381
Qy 5040 ctcaagacagatagttaccgggataaagcgagcagcgttcgggctgaacgggggttcgtgcac 5099

Db 380 CTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTTCGTGCAC 321
Qy 5100 acagcccagcttgagcgaaacgacactacaccgaactgagatacctacacgctgagctatg 5159
Db 320 ACAGCCCAGCTTGGAGCGAAGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATG 261
Qy 5160 agaaagcgcacgcttcccgaaggaggagaaagcgacaggtatccggttaagcggcagggt 5219
Db 260 AGAAAGCGCCACGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGACAGGT 201
Qy 5220 cgaaacagagagcgacagggagcttcacagggggaacgacctgtatctttatagtc 5279
Db 200 CGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGGAAACGCCCTGTATCTTTATAGTCC 141
Qy 5280 tgtcgggtttcgcacacctctgacttgaggtcgatattttgtgatcgtcagggggcg 5339
Db 140 TGTGGGTTTCGCCACCTCTGACTTGAGCGTCTATTTTGTGATGCTCGTCAGGGGGCG 81
Qy 5340 gagcctatgaaaaaacgacgagcaacgagcgccctttttacgggttctcgtctgtgccc 5399
Db 80 GAGCCTATGGAANAACGCCAGCAACGGCGCCTTTTACGGTTCTCTGGCCTTTTGTCTGGCC 21
Qy 5400 ttttgcctca 5408
Db 20 GCAAGCTTA 12
RESULT 12
LOCUS AL044364 954 bp mRNA linear EST 29-FEB-2000
DEFINITION DKFZp434C172_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL044364
VERSION AL044364.1 GI:5432586
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Ansorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansorge W
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
No r1 sequence available.
This clone (DKFZp434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source Location/Qualifiers
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 214 a 273 c 253 g 214 t
ORIGIN
Query Match 13.4%; Score 772.4; DB 9; Length 954;
Best Local Similarity 99.9%; Pred. No. 1.7e-153;

	Matches	773;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	4815	accgctaccagcggtggtttgtttg	ccggatcaagagctacca	aactcttttccg	aaaggt	4874				
DB	954	ACCGCTACCAGCGGTGTTGTTT	GCGGATCAAGAGCTACCA	AACTCTTTTCCG	AAAGGT	895				
QY	4875	aactggcttcagcagcgcgagata	cacaaatactgtcctt	ctagtgtagccg	tagttagg	4934				
DB	894	AACTGGCTTCAGCAGAGCGCAGAT	ACCAATACTGTTCTTCT	AGTGTAGTCGCTAGC	GTAGGTAGG	835				
QY	4935	ccaccacttcaagaactctgtag	caccgcctacatacct	cgctctgcta	atcctgttacc	4994				
DB	834	CCACCACCTCAAGAACTCTGTAG	CACCGCCTACATACTCGT	CTGTGCTAATCTCT	GTACC	775				
QY	4995	agtggctgctgccagtgcgata	aagtcgtctcttacc	gggttgga	ctcaagacgatagtt	5054				
DB	774	AGTGGCTGCTGCCAGTGGCGATA	AAGTCGTGCTTACCG	GGTGGACTCAAGAC	GATAGTT	715				
QY	5055	accggataaaggcgagcggtcg	ggctgaacgggggttc	gtgcacacagcc	agccttgg	5114				
DB	714	ACCGGATAAGGCGCAGCGGT	CGGGCTGAACGGGGG	TTCTGTGCACACAG	CCAGCTTGA	655				
QY	5115	gcgaacgacctacacgcga	aactgagataacct	acagcgtgagct	atgagaaagcgcc	acgt	5174			
DB	654	GCGAACGACCTACACCGAACT	GAGATACCTACAGCGT	GAGTATGAGAAAG	CGCCACGCT	595				
QY	5175	tcccgaggaggaagcgga	caggtatccggta	agcgaggggtcg	gaacagga	gcg	5234			
DB	594	TCCCGAAGGGAGAAAGCGG	ACAGGTATCCGGTA	AGCGGAGGGT	CGGAACAGGAG	CGC	535			
QY	5235	cacgaggagcctccagggg	gaacgcctggtatctt	atagtcctg	tcggttccg	cca	5294			
DB	534	CACGAGGGAGCTTCCAG	GGGAAACGCTGGTATCT	TTATAGTCC	TGTCGGGTTTC	GCCA	475			
QY	5295	cctctgacttgagcgtcg	atttttgtatgctc	gtcagggggcg	gagcctatg	gaaaaa	5354			
DB	474	CCTCTGACTTGAGCGTC	GATTTTGTGATGCT	CGTCGTCAG	GGGGCGGAGCCT	ATGGAAAA	415			
QY	5355	cgccagcaacgcgccctt	tttacgggttcctg	gccttttgc	tggccttttgc	tcacatgtt	5414			
DB	414	CGCCAGCAACGCGGCC	TTTACGGTTCTTGG	CCCTTCTG	GGCCTTTTGCT	CACATGTT	355			
QY	5415	ctttcctgcgttatcc	ctgattctgtggata	aacgtattac	cgccctttg	agtgagctga	5474			
DB	354	CTTTCTCGGTTATCCC	CTGATTCTGTGGATA	AACCGTATTAC	CGCCTTTGAGT	GTAGCTGA	295			
QY	5475	taccgctcgccgcagcc	gaacgagcgcagc	gagtcagtg	agcgaagc	ggaaga	5534			
DB	294	TACCGCTCGCCGCAG	CCGAAACGACCG	GAGCGCAGCG	AGTCACTAG	CGGGAAGA	235			
QY	5535	gcgccc	aatacgcaaacgcgc	ctctccccgcgcgt	tgccgattc	aatgcag	5588			
DB	234	GCGCCAATACGCAAC	CGGCTCTCCCGCGCG	GTGGCGATTCAT	TAAATGCAG	181				

RESULT	13	
BI489087/c		
LOCUS		
DEFINITION		BI489087 990 bp mRNA linear EST 28-AUG-2001 603021122R1 NTH_MGC_114 Homo sapiens cDNA clone IMAGE:5191829 3', mRNA sequence.
ACCESSION		BI489087
VERSION		BI489087.1 GI:15328315
KEYWORDS		EST.
SOURCE		human.
ORGANISM		Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE		1 (bases 1 to 990) NIH-MGC http://mgc.ncl.nih.gov/ .
AUTHORS		National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE		Unpublished (1999)
JOURNAL		
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11479 row: 1 column: 06
 High quality sequence stop: 803.

Location/Qualifiers
 1. .990
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5191829"
 /clone_lib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

BASE COUNT	174 a	316 c	318 g	182 t	ORIGIN
Query Match	13.3%	Score 764.2	DB 10	Length 990	
Best Local Similarity	96.0%	Pred. NO. 9.5e-152			
Matches 869	Conservative 0	Mismatches 28	Indels 8	Gaps 8	
QY 1570	ctgtccccgcgcagccggggcctgttccacagggccgctgctgcagagcgggtgcccccaat	1629			
DB 898	CTCTGTTCCTCCGCCCAACCGGGCACGCTGTCCACAGGCCCTGCTGCAGAACGGTGCCTCCCAAG	839			
QY 1630	ggaccctggggcagcgggtgggcaatgggagagggcccgctgcagggccacgcagctggcccaac	1689			
DB 838	GGG - CCTGGGCCACGTTGGGCCATGGGAAAGGCCCGCTCGCA - GGCCACGCAGCTGGCCAC	781			
QY 1690	cttgtggcgtgtcctccagcggcgccactggtgggaatgaca - cagagctggtagcctgcct	1748			
DB 780	CTTGT - GGCTGTCCTCCA - GCGGCACCTGGTGGGAATGACACCAGAGCTGGTAGCCTGCCT	723			
QY 1749	tcggacacgaccagcgcaggtcctggtgaaccacgaatggcacgtgctgcctcaagaag	1808			
DB 722	TC - GACACGACACGCGCAGGTCTCTGGTGAACACACGAATGGCACGTGCTGCCCTCAAGAAAG	664			
QY 1809	cgtcttcgggttctcttccttcgtgcctggtgtagatggagaccttcctcagtgacacccccaga	1868			
DB 663	CGTCTTCCGGTTCTCTTCGTGCCTGTGGTAGATGGAGACTTCTTCAGTGACACCCCCAGA	604			
QY 1869	ggccctcatcaacgcggggagacctccacggccctgcaggtgctggtgggtggtgaagga	1928			
DB 603	GGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT - TCGTGAAGGA	545			
QY 1929	tgaggcctcgatatttctggtttacggggcccgaggttcagcaaacgagctctct	1988			
DB 544	TGAGGGCTCGTATTTTCTGTGTTTACGGGGCCCCAGGCTTCAGCAAAGACAAACGAGTCTCT	485			
QY 1989	catcagccggccgagttcctggccgggggtgcgggttccccaggttaagtgcct	2048			
DB 484	CATCAGCCGGCTCGAGTCTCTGGCCGGGGTTCGGGGTTCCCCCAGGTAAGTGACCT	425			
QY 2049	ggcagccgaggtgtggtcctgcattacacagactggtgcatccccgagaccggcaag	2108			
DB 424	GGCAGCCGAGGCTGTGGTCTGCATTACACAGACTGGCTGCATCCCGAGGACCCCGGCACG	365			
QY 2109	cctgagggaggccctgagcgaatgtggtggcgaccacaatgtcgtgtcccccggtggccca	2168			
DB 364	CCTGAGGGAGGCCCTTCAGCGCATGTGGTGGGCGACCAACAATGTCTGTGCCCCCGTGGCCCA	305			


```
QY 2169 gctggctggcgactggtgccaggggtgcccgggtctacgacctacgtctttgaacacg 2228
|||||
Db 304 GCTGGCTGGCGACTGGCTGCCAGGGTGCCTGGGTCTACGCCCTACGCTTTGAACACCG 245

QY 2229 tgcttccacgctctcctgccccctgtggatgggggtgccccacggctacgagatcgagtt 2288
|||||
Db 244 TGCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGGTGCCCCACGGCTACGAGATCGAGTT 185

QY 2289 catctttggatccccctggacccctctgaaactacacggcagagagagaaaaatcttcgc 2348
|||||
Db 184 CATCTTTGGGAT-CCCCTGGACCCCTCTGAAACTACACGGCAGAGGAGAAATCTTCGC 126

QY 2349 ccagcgactgatcgatactgggccaaactttgccgcacagggggtacccaatgagccccg 2408
|||||
Db 125 CCAGCGACTGATCGGATACTGGGCCAACTTGCCCCGCACAGGGGATCCCAATGAGCCCCG 66

QY 2409 agaccccaagggccccacaaatggccccctacacggcggggtcagcagtagtct 2468
|||||
Db 65 AGACCCCAAGGGCCCCACAATGGCCCCCTACACGGCGGGGGCTCAGCAGTACGTTAGTCT 6

QY 2469 ggacc 2473
|||||
Db 5 GGACC 1

RESULT 14
BI488594
LOCUS 603021122F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191829 5',
DEFINITION mRNA sequence.
ACCESSION BI488594
VERSION BI488594.1 GI:15327822
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 931)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11479 row: 1 column: 06
High quality sequence stop: 771.
Location/Qualifiers
1. .931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5191829"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dr
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."
```

BASE COUNT 152 a 294 c 306 g 179 t

ORIGIN

Query Match

13.0%; Score 750.4; DB 10; Length 931;

RESULT 15

BG923768/c

LOCUS

DEFINITION

BG923768

602825893F1

mRNA sequence.

832 bp

mRNA

linear

EST 05-JUN-2001

IMAGE:4954486 5',

```
Best Local Similarity 96.9%; Pred. No. 8e-149;
Matches 882; Conservative 0; Mismatches 16; Indels 12; Gaps 11;

QY 967 cgtggggccggctgcgggcatcgccttcgcttcgcttcctctctctctctctctctctc 1026
|||||
Db 1 CGTGGGGCCGCTGCGGGGCATTCGCCTGAAGACCCCGGGGCGCTCTCTGCTTTC 60

QY 1027 ctgggcatcccccttgcggagccacccttcgggacccctgcgcttctgccacggagccc 1086
|||||
Db 61 CTGGGCATCCCTTTCGGAGCCACCCATGGACCCCGCTCGCTTTCGCCACGGAGCCC 120

QY 1087 aagcagccttggtcaggggtggtagacgctacaaccttcagagatgtctgtaccacatat 1146
|||||
Db 121 AAGCAGCCTTGTGAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGTACCAATAT 180

QY 1147 gtggacacccctataccacagggttttgagggcaccgagatgtggaaccccaaccgtgagctg 1206
|||||
Db 181 GTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAACCGTGAGCTG 240

QY 1207 agcaggactgcctgtacctaactgaactgttgacacacataccccccggtcacatccccacc 1266
|||||
Db 241 AGCGAGGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGCTACATCCCCCACC 300

QY 1267 cctgtcctcgtctggtatggtgggtggtgtctctacagtggggcctcctccttg-gacgt 1325
|||||
Db 301 CCTGTCTCCTGTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCCTCCTTGTGACGT 360

QY 1326 gtacgatggccgcttcttggtacagggccgagagactgtgctggtgtccatgaactaccg 1385
|||||
Db 361 GTACGATGGCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATGAACCTACCG 420

QY 1386 ggtgggagccttctgctcctgcccctgcgggggagccgagagcccggggcaatgtggg 1445
|||||
Db 421 GGTGGGAGCCTTTGGCTTCTTGGCCCTGCGGGGAGCCGAGAGGCCCGGGCAATGTGGG 480

QY 1446 tctcctggatcagaggctggccctgcagtggtggtgcaggagaacgtggcagccttcggggg 1505
|||||
Db 481 TCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCCTTCGGGGG 540

QY 1506 tgacccgacatcagtgacgctgtttgggagagcgcgggagccgctcggtgggcatgca 1565
|||||
Db 541 TGACCCCGACATCAGTACGCTGTTTGGGAGAGCGCGGGAGCCGCTCGGTGGCATGCA 600

QY 1566 cctgctgtccccgccagccggg-gcctgtccacagggccgtgtgcagagc-ggtgcc 1623
|||||
Db 601 CCTGCTGTCCCCGCCAGCCGGGGCCCTGTCCACAGGGCCGTGCTGCAGAGCGGGTGCC 660

QY 1624 cccaatggaccctggccacggtgggcatgggagagggccgctgcagggccac--gcagc 1681
|||||
Db 661 CCCAATGGACCCCTGGGCCACCGTGGGCATGGGAGAGGCCCGCTCGCAGGGCCACAGCAGCT 720

QY 1682 tggccccacctgtgtgggctgt-gcctcagggcgccactggtgggaatgacacagagct-ggt 1739
|||||
Db 721 TGGCCCCACCTTGTGGGCTGTCCCTCCAGCGGGCAGCTGGTGGGAATGACACAGAGCTGGGT 780

QY 1740 agcctgcc-ttcggacacgacc-agcgaaggtcctggtgaaccacgaatggcacgtgctg 1797
|||||
Db 781 AGCCTGCCTTTCGGACACGACCAAGCGCAGGTCTCTGGTGAACACACGAATGGCAGCTGCTG 840

QY 1798 cctcaagaagcgtcttccggttctcctcgtgcctggtggtagatgagagacttcctcagt 1857
|||||
Db 841 -CTCAAGAAAGCGTCTTCCGG-TCTCCTTCGTGCCTGT-GTCCATTGCGACTTCCTCAGG 897

QY 1858 gacacccacag 1867
| || |
Db 898 AAACCCACAGAG 907
```


ACCESSION BG923768
VERSION BG923768.1 GI:14304244
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 832)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes Of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10915 row: j column: 23
High quality sequence start: 32
High quality sequence stop: 823.
FEATURES
source
1..832
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4954486"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: Sali;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 201 a 209 c 196 g 226 t
ORIGIN

Query Match 12.9%; Score 743.6; DB 10; Length 832;
Best Local Similarity 99.2%; Pred. No. 2.2e-147;
Matches 768; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 3904 ttctgccccgaagaacggttttccaatgatgagcacttttaaaagtctgtatgtggcg 3963
|||||
Db 805 TTTTCGCCCCGAAGAACGTTTCCCAATGATGAGCAC-TTTAAAGTTCTGCTATGTGGCGG 747
QY 3964 gtattatcccgattattgacggcggaagagcaactcgg-tcgccgcatacacatttctca 4022
|||||
Db 746 GTATTATCCCGTATTGACGCGCGGCAAGAGCAACTCGGTTGCGCGCATACACTATTCTCA 687
QY 4023 gaatgacttggttgagtactaccagtcacagaaaagcatcttacggatggcatgacagt 4082
|||||
Db 686 GAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGT 627
QY 4083 aagagaattatgcagtgctgccataaccatgagtataaacactgcggccaacttacttct 4142
|||||
Db 626 AAGAGAATTATGCAGTGTCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCT 567
QY 4143 gacaacgatcgaggagaccgaaggagctaaccgcctttttgcacaacatgggggatcatgt 4202
|||||
Db 566 GACAAACGATCGGAGGACCGAAGAGCTAACCGCTTTTTCACAACATGGGGGATCATGT 507
QY 4203 aactgccttgatcggttgggaaccggagctgaatgaagccaactattaaactgagcgctga 4262
|||||
Db 506 AACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGA 447
QY 4263 caccagatgcctgtagcaatggcaacaacgttgcgcaactattaaactggcggaactact 4322
|||||
Db 446 CACCACGATGCCTGTAGCAATGSCAACCAACGTTGCGCAAACTATTAACTGGCGAACTACT 387

QY 4323 tactctagcttccccggcaacaattaatagactggatggaggcggaataaaagttgcaggacc 4382
|||||
Db 386 TACTCTAGCTTCCCGGCAACAATAATAGACTGGATGGAGCGGATAAAAGTTGCAGGACC 327
QY 4383 actctgcgctcgcccccttcggctggctggtttattgtgataaaatctggagccggtga 4442
|||||
Db 326 ACTTCTGCGCTCGGCCCTTCCGGCTGGTGGTTTATTGCTGATAAAATCTGGAGCCGGTGA 267
QY 4443 gcgtgggtctcgcggtatcatctgcagcactggggccagatggtaagccctcccgtatcgt 4502
|||||
Db 266 GCCTGGGTCTCGCGGTATCATTCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGT 207
QY 4503 agttatctacacgacgaggggagtcaggcaactatggatgaacgaataagacagatcgctga 4562
|||||
Db 206 AGTTATCTACACGACGCGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGA 147
QY 4563 gataggtcctcactgattaaagcattggtaactgtcagaccaagtttactcatatatact 4622
|||||
Db 146 GATAGGTGCCTCACTGATTAAAGCATTTGTTAACTGTCAGACCACCAAGTTTACTCATATATACT 87
QY 4623 ttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatcct 4676
|||||
Db 86 TTAGATTGATTTAAACCTCATTTTAAATTTAAAGGATCTCGCCGAAGATCCT 33

Search completed: August 31, 2002, 15:25:56
Job time: 6459 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:27:40 ; Search time 153.1 Seconds
(without alignments)
9252.570 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcatgcctgcaggtc.....ctatgaccatgattacgcca 5767

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2637.8	45.7	8854	3 US-09-053-549-1	Sequence 1, Appl
C 2	2632.8	45.7	7387	4 US-09-238-356-28	Sequence 28, Appl
C 3	2631.8	45.6	4283	1 US-08-343-401A-3	Sequence 3, Appl
C 4	2631.8	45.6	4283	1 US-08-445-265A-1	Sequence 1, Appl
C 5	2631.8	45.6	4283	3 US-08-990-442-1	Sequence 1, Appl
C 6	2630	45.6	6295	2 US-08-659-206A-4	Sequence 4, Appl
7	2628.2	45.6	4045	4 US-08-464-700-54	Sequence 54, Appl
C 8	2628	45.6	4950	3 US-08-789-333F-58	Sequence 58, Appl
C 9	2628	45.6	9639	4 US-09-147-208-26	Sequence 26, Appl
C 10	2627	45.6	7566	2 US-08-232-016-23	Sequence 23, Appl
C 11	2627	45.6	7639	2 US-08-232-016-22	Sequence 22, Appl
12	2624.8	45.5	4713	4 US-09-194-285-7	Sequence 7, Appl
13	2624.8	45.5	4724	4 US-09-194-285-8	Sequence 8, Appl
14	2623.2	45.5	7560	4 US-08-844-274-20	Sequence 20, Appl
15	2620.8	45.4	6350	2 US-08-383-335A-8	Sequence 8, Appl
C 16	2616	45.4	11284	3 US-08-978-741-5	Sequence 5, Appl
17	2579.8	44.7	3875	4 US-09-039-982A-21	Sequence 21, Appl
18	2579.8	44.7	3875	4 US-09-039-641-21	Sequence 21, Appl
19	2579.8	44.7	3875	4 US-09-039-762A-21	Sequence 21, Appl
20	2579.8	44.7	3878	4 US-09-039-982A-27	Sequence 27, Appl
21	2579.8	44.7	3878	4 US-09-039-641-27	Sequence 27, Appl
22	2579.8	44.7	3878	4 US-09-039-762A-27	Sequence 27, Appl
23	2579.8	44.7	3883	4 US-09-039-982A-30	Sequence 30, Appl
24	2579.8	44.7	3883	4 US-09-039-641-30	Sequence 30, Appl
25	2579.8	44.7	3883	4 US-09-039-762A-30	Sequence 30, Appl
26	2579.8	44.7	3908	4 US-09-039-982A-24	Sequence 24, Appl
27	2579.8	44.7	3908	4 US-09-039-641-24	Sequence 24, Appl

28	2579.8	44.7	3908	4 US-09-039-762A-24	Sequence 24, Appl
29	2575.8	44.7	6111	4 US-09-538-414-9	Sequence 9, Appl
30	2545.6	44.1	3822	3 US-08-675-566-8	Sequence 8, Appl
31	2545.6	44.1	3861	3 US-08-675-566-11	Sequence 11, Appl
32	2545.6	44.1	3888	3 US-08-675-566-12	Sequence 12, Appl
33	2545.6	44.1	3955	3 US-08-675-566-10	Sequence 10, Appl
34	2545.6	44.1	4009	3 US-08-675-566-9	Sequence 9, Appl
35	2545.6	44.1	4503	3 US-08-675-566-7	Sequence 7, Appl
36	2544.6	44.1	5067	3 US-08-675-566-20	Sequence 20, Appl
37	2544.6	44.1	5147	3 US-08-675-566-24	Sequence 24, Appl
38	2544.6	44.1	5241	3 US-08-675-566-23	Sequence 23, Appl
39	2451.6	42.5	10306	3 US-08-716-351A-4	Sequence 4, Appl
40	2451.6	42.5	10970	3 US-08-716-351A-5	Sequence 5, Appl
C 41	2449.4	42.5	5639	3 US-09-175-690A-1	Sequence 1, Appl
C 42	2449.4	42.5	6306	5 PCT-US94-00558-1	Sequence 1, Appl
C 43	2437.4	42.3	4800	5 PCT-US94-07779-1	Sequence 1, Appl
C 44	2418	41.9	8225	4 US-08-793-618-1	Sequence 1, Appl
45	2416.4	41.9	19307	3 US-08-836-022A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-053-549-1
; Sequence 1, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..3694
; OTHER INFORMATION: /product= "hyFLIB protein"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8854
; OTHER INFORMATION: /note= "Sequence of pCIB5520
; Patent No. 6121521
; OTHER INFORMATION: containing coding sequence for hyFLIB protein"
US-09-053-549-1

Query Match		45.7%;	Score 2637.8;	DB 3;	Length 8854;
Best Local Similarity		99.9%;	Pred. NO. 0;		
Matches 2639;	Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
QY	3127	gaagcgaattcactggcgcgtgcttttacaaactgctgactgggaaacccctggcgttacc	3186		
Db	3888	GTACCGAATTCACTGGCCGTCGTTTACAACTGCTGACTGGGAAACCCCTGGCGTTACC	3947		
QY	3187	caacttaatgccttgcagcacatcccccttgcgagctggcgtgtaataagcgaagagccc	3246		
Db	3948	CAACTTAATCGCCTTGACGACACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCC	4007		
QY	3247	cgcaccgatgcgccttcccaacagttgcgagccctgaatggcgaatggcgccctgatcgcg	3306		
Db	4008	CGACCGATCGCCCTTCCCAACAGTTGGCAGCCTGAATGGCGAATGGCGCCTGATCGCG	4067		
QY	3307	tattttctccttacgcattctgctgctggttatttcacaccgcgcatatggtgcactctcagtaca	3366		
Db	4068	TATTTTCTCCTTACGCATCTGTGCGGTATTTTCACACCGCATATGGTGCATCTCAGTACA	4127		
QY	3367	atctgctctgatgcgcgcatagtttaagccagcccccgacaccccccaacaccccgctgacgcg	3426		
Db	4128	ATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCCGCCNACACCCCGCTGACCGG	4187		
QY	3427	ccctgacgggcttctgtctcccgccatccgcttacagacaagctgtgacctctccggg	3486		
Db	4188	CCCTGACGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGG	4247		
QY	3487	agctgcatgtgtcagagggttttccaccgtcatcacgaaacgcgcgagacgaaaggccctc	3546		
Db	4248	AGCTGCATGTGTCAAGGTTTTCACCGTCATCACCGAAACGCGGAGACGAAAGSGCCTC	4307		
QY	3547	gtgatacgcctattttataggttaatgtcatgataataatggtttcttagacgtcaggt	3606		
Db	4308	GTGATACGCCCTATTTTATAGGTTAATGTCAATGATAATAATGGTTTCTTAGACGTCAGGT	4367		
QY	3607	ggcacttttcggggaaatgtgcggaacccctatttggttatttttctctaaatacattca	3666		
Db	4368	GGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTGTGTTATTTTCTAAATACATTCA	4427		
QY	3667	aatatgtatccgctcatgagacaataaaccttgataaatgcttcaataatattgaaaaagg	3726		
Db	4428	AATATGTATCCGCTCATGAGACAATAAACCTTGATAAATGCTTCAATAATATTGAAAAAGG	4487		
QY	3727	aagagtatgagtattcaacatttcccggtgcgccttattcccttttttgcggcaatttgc	3786		
Db	4488	AAGAGTATGAGTATTCACATTTCCGTGTCGCCCTTATTCCTTTTTCGGGCATTTTGC	4547		
QY	3787	cttccctgttttgcctccagaaaacgctggtgaaagtaaaagatgctgaagatcagttg	3846		
Db	4548	CTTCTGTTTTTGTCTACCCAGAAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTG	4607		
QY	3847	ggtgcacgagtgggttacatcgaactggatctcaacagcggtgaagatccttgagagtitt	3906		
Db	4608	GGTGACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTT	4667		
QY	3907	cgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtta	3966		
Db	4668	CGCCCCGAAGAACGTTTCCAATGATGAGCACITTTTAAAGTTCTGCTATGTGGCGCGGTA	4727		
QY	3967	ttatcccgattatgacgcggggaagagcaactcggtcgccgcgcatacactattctcagaat	4026		
Db	4728	TTATCCCGTATTGACCGCGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAA	4787		
QY	4027	gacttgggtgactaccagtcacagaaaagcatcttacggatggcgatgacagtaaga	4086		
Db	4788	GACTTGGTTGAGTACTCACCACTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGA	4847		
QY	4087	gaattatgcagtgcgcataaaccatgagtgaataacactgcggccaacttactctgaca	4146		
Db	4848	GAATTATGCAGTGTGCCATAAACCATGAGTGATAACACTGCGGCCAACCTTACTTCTGACA	4907		

QY	4147	acgatcggaggaccgaagagctaaaccgctttttgcacaacatgggggatcatgttaact	4206
Db	4908	ACGATCGGAGACCGAAGGAGCTAAACCGCTTTTTCACAACATGGGGATCATCTAACT	4967
QY	4207	cgccttgatcgttgggaacccggagctgaatgaagccataccaaacacgacgagcgtgacacc	4266
Db	4968	CGCCTTGATCGTTGGGAACCCGAGCTGAATGAAGCCATACCAACACGACGCGTGACACC	5027
QY	4267	acgatgcctgtagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactactact	4326
Db	5028	ACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTACT	5087
QY	4327	ctagcttccccggcacaacaaattaatagactggatggagcggaataaaagtgcaggacacatt	4386
Db	5088	CTAGCTTCCCGCAACAATAATAGACTGGATGGAGCGGATAAAAGTTGCAGGACCACCT	5147
QY	4387	ctgcgctcgcccttccgctggctggtttatttgcgtgataaaactctggagccggtgagcgt	4446
Db	5148	CTGCGCTCGGCCCTTCCGCGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGCTGAGCGT	5207
QY	4447	gggtctcgcggtatcatctgcagcactggggccagatgggtaaagccctccccgtatcgtagtt	4506
Db	5208	GGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGSTAAAGCCCTCCCGTATCGTAGTT	5267
QY	4507	atctacacgacggggagtcaggcaactatggatgaacgaataagacagatcgcgtgagata	4566
Db	5268	ATCTACACGACGGGGAGTCAAGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATA	5327
QY	4567	ggtgcctcactgattaaagcattggtaactgtcagacacaaagtttactcatatatactttag	4626
Db	5328	GGTGCCTCACTGATTAAGCATTTGGTAACCTGTAGACACCAAGTTTACTCATATATACTTAG	5387
QY	4627	attgatttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttttgataaat	4686
Db	5388	ATTGATTTAAAACTTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAAT	5447
QY	4687	ctcatgaccaaaaatcccttaacgtgagtttcttcctccactgagcgtcagacccccctagaa	4746
Db	5448	CTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAA	5507
QY	4747	aagatcaaaaggatcttcttgagatcccttttctgcggttaactctgctgcttgcaaaaca	4806
Db	5508	AAGATCAAAAGGATCTTCTTGAGATCCTTTTTCGCGGTAAATCTGCTGCTTGCAAAACA	5567
QY	4807	aaaaaacaccgcctaccagcggtgttgttttgcgggatacaagactaccaactctttt	4866
Db	5568	AAAAAACCCACCGCTACCGCGGTGGTTGTTTGCCGGATCAAGAGCTACCAACTCTTTT	5627
QY	4867	ccgaaggtaactggttcagcagagcgcagataaccaataactgtccctctctagttagccg	4926
Db	5628	CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTTAGTGTAGCCG	5687
QY	4927	tagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctctgctaact	4986
Db	5688	TAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATC	5747
QY	4987	ctgttaccagtggtgctgcccagtggcgataagtcgtcttaccgggttggaactcaaga	5046
Db	5748	CTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTCTTACCGGGTTGGACTCAAGA	5807
QY	5047	cgatagttaccgggataaaggcgcgagcggtcgggctgaacggggggttcgtgcacacagccc	5106
Db	5808	CGATAGTTACCGGATAAGGCGCAGCGGTGCGGGCTGAACGGGGGGTTCGTGCCACACAGCCC	5867
QY	5107	agcttggagcggaacgacctacaccgaactgagataacctacagcgtgagctatgagaaagc	5166
Db	5868	AGCTTGGAGCGGAACGACCTACACCCGAACCTGAGATACCTACAGCGTGAGCTATGAGAAAGC	5927
QY	5167	gccacgcttcccgaaaggagaaaggcgagaggtatccgggttaagcggcagggttcggaaca	5226
Db	5928	GCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTTCGGAACA	5987

QY 4204 actgccttgatcgttgggaaccggagctgaatgaagccataccaaacgagcagcgtgac 4263
Db 6285 ACTGCCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGAC 6226
QY 4264 accacgatgcctgtagcaaatggcaacaacgcttgcgcaaaactattaaactggggaactactt 4323
Db 6225 ACCACGATGCCGTGATAGCAATGCCAACACGTTGCCGCAAACTATTAACTGGCGAACTACTT 6166
QY 4324 actctagcttcccggaacaataataagactggatggaggcgataaaagtgcaggacca 4383
Db 6165 ACTCTAGCTTCCCGCAACAATAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCA 6106
QY 4384 cttctgcgctcggcccttccggctggctggtttattgctgataaaatctggagccggtgag 4443
Db 6105 CTCTCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAG 6046
QY 4444 cgtgggtctcgcggtatcatctgcagcactggggccagatggttaagccctcccgctatcgta 4503
Db 6045 CGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTA 5986
QY 4504 gttatctacacgagcggggagtcaggcaactatggatgaacgaaaaatagacagatcgcgtgag 4563
Db 5985 GTTATCTACACGACGGGGAGTCAGGCCAACTATGGATGAACGAAATAGACAGATCGCTGAG 5926
QY 4564 ataggtgcctcactgattaaagcattggttaactgctcagaccaaagtcttactatatactt 4623
Db 5925 ATAGGTGCCTCACTGATTAAGCATTTGGTAACCTGTCAGACCAAGTTTACTCATATATACTT 5866
QY 4624 tagattgatttaaaacttcaatttttaatttaaaaggatctaggtgaagatcctttttgat 4683
Db 5865 TAGATTGATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTAT 5806
QY 4684 aatctcatgaccaaatacccttaacgtgagtttctgctccactgagcgtcagacccccgta 4743
Db 5805 AATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCCGTA 5746
QY 4744 gaaaagatcaaaaggatcttcttgagatcctttttctgcgcgtaactcgtcgttgcaa 4803
Db 5745 GAAAGATCAAAAGGATCTTCTTGAGATCCCTTTTCTGCGGTAATCTGCTGTGCAA 5686
QY 4804 acaaaaaaaccccgctaccagcgggtggttctgctgcggatcaagagactaccaactctt 4863
Db 5685 ACAAAAAAACCCCGCTACCAGCGGTGGTTTGTTCGGGATCAAGAGCTACCAACTCTT 5626
QY 4864 ttccggaaggtaactggcttcagcagagcgcagataccaaataactgtccttctagtgtag 4923
Db 5625 TTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCGAGATACCAATACTGTCTCTTAGTGTAG 5566
QY 4924 ccgtagttaggccaccacttcaagaactctgtagccgcctacataacctcgtctgcta 4983
Db 5565 CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTA 5506
QY 4984 atcctgttacagtggtgctgctccagtgccagtgccggtgctgaacgggggttctgactca 5043
Db 5505 ATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTGTCTTACCGGTTGGACTCA 5446
QY 5044 agacgatagttaccgggataaggcgccagcgggtgctgaacgggggttctgctgcacacag 5103
Db 5445 AGACGATAGTTACCGGATAAGCGCGACGCGTCCGGCTGAACGGGGGGTTCGTGCACACAG 5386
QY 5104 ccagcttggagcgaacgacctacacccgaactgagatacctacagcgtgagctatgagaa 5163
Db 5385 CCCAGCTTGGAGCGCAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA 5326
QY 5164 agcggccacgcttcccgaaaggagaaaggcgagaggtatccggtaagcggcagggctcgga 5223
Db 5325 AGCGCCACGCTTCCCGAAGGGAGAAAGGGCGACAGGTATCCGGTAAGCGCGAGGGTCGGA 5266
QY 5224 acaggagagcgcagggagcttccaggggaacgcctggtatctttatagtcctgctc 5283
Db 5265 ACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCTCTGTC 5206
QY 5284 ggggttcgccacctctgacttgagcgtcgatttttgtgatgctcagggggggcgagc 5343

Db 5205 GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGC 5146
QY 5344 ctatggaaaaacgcagcaacgcggccttttttagcggttcctggttcttgccttttgcctttt 5403
Db 5145 CTATCGAAAAACGCCAGCAACGGGCCCTTTTACGGTTCTTGCCCTTTTGTGGCCTTTT 5086
QY 5404 gctcacatgttcttctcgtgttatccctcgtattcctcgttggtgataaacogtattaccgccttt 5463
Db 5085 GCTCACATGTTCTTCTGCTGCTTATCCCTGATTTCTGTGGATAACCGTATTACGCCCTTT 5026
QY 5464 gagtgcgtgatacgcgtcgcgcagccgaacgcagcagcagcagcagcagcagcagcag 5523
Db 5025 GAGTGAGCTGATACCGCTCGCCGACGCGAAGCGGAGCGAGCGAGTCAAGCGGAG 4966
QY 5524 gaagcgggaagcgcgcaataacgcaaacccgcctctcccgcgcttggccgatttcattaa 5583
Db 4965 GAAGCGGAAGAGCGCGCAATACGCAACCGCCTCTCCCGCGCTTGCCCGATTTCATTAA 4906
QY 5584 tgcagctggcacgacaggttctccgactggaaagggggcagtgagcgcgaacgaattaat 5643
Db 4905 TGCAGCTGGCACGACAGGTTTCCCGACTCGAAAGCGGCGCAGTGAGCGCAACGAATTAAT 4846
QY 5644 gtgagttagctcactcattaggtcagccccaggctttacacatttatgcttccggctcgtatg 5703
Db 4845 GTGAGTTAGCTCACTCATTAGGCACCCCGGCTTACACTTTATGCTTCCGGCTCGTATG 4786
QY 5704 ttgtgtggaattgtgagcgggataaacaatttcacacaggaacagcgtatgacctgattac 5763
Db 4785 TTGTGTGGAATTGTGAGGGGATAACAATTTTCACACAGGAACAGCTATGACCATGATTAC 4726
QY 5764 gccca 5767
Db 4725 GCCA 4722

RESULT 3

US-08-343-401A-3/c
; Sequence 3, Application US/08343401A
; Patent No. 5661132
; GENERAL INFORMATION:
; APPLICANT: Swain, William F
; APPLICANT: Macklin, Michael D
; APPLICANT: Eriksson, Elof
; APPLICANT: Andree, Christophe
; TITLE OF INVENTION: Improved Wound Healing
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: PO Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/343,401A
; FILING DATE: 22-NOV-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 11-229-9103-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: pWRG1630
; FEATURE:
; NAME/KEY: exon
; LOCATION: 713..721
; FEATURE:
; NAME/KEY: exon
; LOCATION: 981..1253
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1253)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 713..1049
US-08-343-401A-3

Query Match 45.6%; Score 2631.8; DB 1; Length 4283;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	3133	aattcaactggcgcgtcgttttacaaacgtcgtgactgggaaacccctggcgttacccaaactt	3192
Db	4265	AATTCACTGGCCGTCGTTTTACAAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAACTT	4206
Qy	3193	aatcgccctgcagcacatcccccttcgcgcagctggcgtaatagcgaagagcccgacc	3252
Db	4205	AATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCGCAC	4146
Qy	3253	gatcgcccttcccaacagttgcgcgcgcctgaatggcgaaatggcgctgatcggtatttt	3312
Db	4145	GATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGCGCTGATGCGGTATTTT	4086
Qy	3313	ctccttaacgcatctgtgcggtatttcacacccgcataatggtgcactctcagtaacaatctgc	3372
Db	4085	CTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACCTCTCAGTACAATCTGC	4026
Qy	3373	tctgatccgcgcatagttaagccagcccgacaccccgcaacacccgcctgacgcgcctga	3432
Db	4025	TCTGATGCCGCATAGTTAAGCCAGCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA	3966
Qy	3433	cgggctgtctgtctcccggcatccgcttacagacaagctgtgacgcgtctccggagctgc	3492
Db	3965	CGGGCTGTCTGTCTCCGGCATCCGGTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC	3906
Qy	3493	atgtgtcagaggttttcaccgttcatacccgaaacgcgcgagacgaaagggcctcgtgata	3552
Db	3905	ATGTGTcAGAGGTTTTACCCGTTCATCACCgAAACGCGGAGACGAAAGGGCCTCGTGATA	3846
Qy	3553	cgcctattttataggttaatgctcatgataataatgggtttcttagacgtcaggtggcact	3612
Db	3845	CGCCTATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACT	3786
Qy	3613	tttcggggaaatgtgcgcgggaacccctatttgttttatttttctaaataacattcaaatatg	3672
Db	3785	TTTTCGGGAAATGTGCGGGAAACCCCTATTGTGTTATTTTCTTAAATACATTCAAATATG	3726
Qy	3673	tatccgcctcatgagacaaataacccctgataaaatgcttcaataataattgaaaaaggaagagt	3732
Db	3725	TATCCGCTCATGAGACAAATAACCCCTGATAAAATGCTTCAATAATATTGAAAAAGGAAGAGT	3666
Qy	3733	atgagtattcaacatttcctcgtgcgccttattcccttttttgcggcattttgccttccct	3792
Db	3665	ATGAGFATTCAAACATTTCCGCTGTCGCCCTTATTCCTTTTTTTCGGCATTTTGCCTTCCT	3606
Qy	3793	gtttttgctcaccagaaacgctggtgaaagtataaagatgctgaagatcagttgggtgca	3852
Db	3605	GTTTTTGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCA	3546

Qy	3853	cgagtgggttacatcgaaactggatctcaacagcggtaagatcccttgagaggttttcgcccc	3912
Db	3545	CGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC	3486
Qy	3913	gaagaacgttttccaatgatgagcacttttaaagttctctctatgtggcgcggtattatcc	3972
Db	3485	GAAGAACGTTTCCAATGATGAGCACITTTAAAGTTCTGCTATGTGGCGGGTATTATCC	3426
Qy	3973	cgtattgacgcggggaagagcaactcgtgcgcgcgcatacacactattctcagaatgacttg	4032
Db	3425	CGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAAATGACTTG	3366
Qy	4033	gttgagtactcaaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaatta	4092
Db	3365	GTTGAGTACTCACCAAGTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA	3306
Qy	4093	tgacgtgctgccataaaccatgagtataaacactcggccaaacttactcttgacaacgac	4152
Db	3305	TGCAGTGTGCCATAAACCATGAGTGATAACACTGCGGCCAACITTACTTCTGACAACGATC	3246
Qy	4153	ggaggaccgaaggagctaaaccgcttttttgcaacaacatgggggacatcgttaactcgcctt	4212
Db	3245	GGAGGACCGAAGGAGCTAACCGCTTTTTCACACAACATGGGGGATCATGTAACTCGCCTT	3186
Qy	4213	gatcgttgggaaccggagctgaatgaagcacaataccaaaacgacgagcgtgacacccacgatg	4272
Db	3185	GATCGTTGGAAACCGGAGCTGAATGAAGCATACCAAAACGACGAGCGTGACACCACGATG	3126
Qy	4273	cctgtagcaatggcaacaaacgttgcgcaaaactattaaactggcgaactacttactctagct	4332
Db	3125	CCTGTAGCAATGCCAAACAAACGTTTGCgAAACTATTAACTGGCGAACTACTTACTCTAGCT	3066
Qy	4333	tcccggcaacaataatagactggatggagggcgggataaaagtgcaggaccactctcgcgc	4392
Db	3065	TCCCGGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACTTCTCGGC	3006
Qy	4393	tcggcccttcgggctggctggtttattgtctgataaaatctggagccggtgagcgtgggtct	4452
Db	3005	TCGGCCCTTCGGCTGGCTGGTTATTGCTGTATAAATCTGGAGCCGCTGAGCGTGGGTCT	2946
Qy	4453	cgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagtattctac	4512
Db	2945	CGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC	2886
Qy	4513	acgacggggagtcaggcaactatggatgaacgaaatagacagatcgcgtgagataggtgcc	4572
Db	2885	ACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC	2826
Qy	4573	tcactgattaagcattggttaactgtcagaccaagtttactcatatataactttagattgat	4632
Db	2825	TCACTGATTAAAGCATTTGGTAACGTGTcAGACCAAGTTTACTCATATATACTTTAGATTGAT	2766
Qy	4633	ttaaaacttcatttttaatttaaaggatctaggtgaagatccctttttgataaactctcatg	4692
Db	2765	TTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCATG	2706
Qy	4693	accaaaaatcccttaacgtgagtttttcgttccactgagcgtcagaccccgtagaaaaagatc	4752
Db	2705	ACCAAAATCCCTTAACGTGAGTTTTCGTTCCTACTGAGCGTCAGACCCCGTAGAAAAGATC	2646
Qy	4753	aaaggatcttcttgagatccctttttttctgcgcgtaactctgctgcttgcaaaaaaaaa	4812
Db	2645	AAAGGATCTTCTTGAGATCCCTTTTTTCTGGCGGTAATCTGCTGCTTGCAAAACAAAAAA	2586
Qy	4813	ccaccgctaccagcgggtggtttgttgcgggatcaagagctaccaactcttttccgaag	4872
Db	2585	CCACCGCTACCAGCGGTGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAG	2526
Qy	4873	gtaactggcttcagcagcgcagatatacaaaaactgtccttcttagtgcgtagcctagtta	4932
Db	2525	GTAACGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTTCTTAGCTAGCCGTAGTTA	2466
Qy	4933	ggccaccacttcaagaactctgtagcacgcgcctacataacctgcctctgctaactcctgta	4992

||||| 3665 ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTGTGGGCATTTGCGCTTCCT 3606
QY 3793 gtcttctgctcaaccagaaacgcctggtgaaagtataaagatgctgaaagatcaagtgtggtgca 3852
Db 3605 GTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCA 3546
QY 3853 cgagtgggttacatcgaaactggatctcaacagcggtaagatcccttgagagttttcgcccc 3912
Db 3545 CGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC 3486
QY 3913 gaagaacgttttccaatgatgagcaccttttaaagtctgctatgtggcggtattatcc 3972
Db 3485 GAAGNACGTTTCCAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGCGGTATTATCC 3426
QY 3973 cgtattgacgcgggcaagagcaactcgtgcgcgcgcatacacactattctcagaatgacttg 4032
Db 3425 CGTATTGACGCCGGCAAGAGCAACTCGSTCGCCGCATACACTATTCTCAGAATGACTTG 3366
QY 4033 gttgagtactaccagtcacagaaaaagcatcttacggatggcatgacagtaagaaatta 4092
Db 3365 GTTGAGTACTACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAAATTA 3306
QY 4093 tgcagtgctgcataaaccatgagtgataaaacactgcggccaaacttacttctgacaaacgatc 4152
Db 3305 TGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATC 3246
QY 4153 ggaggaccgaaggagctaaccgcttttttgacaacaatggggatcatataactcgctt 4212
Db 3245 GGAGGACCGAAGAGCTAACCGCTTTTGTGACAAACATGGGGATCATCTAACTCGCCTT 3186
QY 4213 gatcgttgggaacgggagctgaatgaagccataccaaacgacgagcgtgacaccacgatg 4272
Db 3185 GATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGATG 3126
QY 4273 cctgtagcaatggcaacaacgcttgcgcaaaactattaaactggcgaactacttactctagct 4332
Db 3125 CCTGTAGCAATGCCAACACACGTTGCGCAAACTATTAACTGGCGAACTACTTACTTAGCT 3066
QY 4333 tcccgcccttcggcgtgctggttttatgtcgataaaactctggagccggtgagcgtggtct 4392
Db 3065 TCCCGGCNACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACTTCTGCGC 3006
QY 4393 tcggcccttcggcgtgctggttttatgtcgataaaactctggagccggtgagcgtggtct 4452
Db 3005 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCCGGTGAGCGTGGGTCT 2946
QY 4453 cgcggtatcatgacgacactggggccagatggttaagccctcccgatctagttatctac 4512
Db 2945 CGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCCGTATCTATCTAC 2886
QY 4513 acgacggggagtcaggcaactatggatgaacgaaaaatagacagatcgcctgagataggc 4572
Db 2885 ACGACGGGGAGTCAGGCAACTATGGATGAACGAAAATAGACAGATCGCTGAGATAGGTGCC 2826
QY 4573 tcactgattaaagcatggtaactgtcagaccaagtttactcatatataacttttagattgat 4632
Db 2825 TCACTGATTAAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT 2766
QY 4633 ttaaaacttcatttttaatttaaaaggatctagggtgaagatccctttttgataaatctcatg 4692
Db 2765 TTAAACTTCAATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATG 2706
QY 4693 accaaaaatcccttaacgtgagtttcttcactgagcgtcagacccccctagaaaaagatc 4752
Db 2705 ACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATC 2646
QY 4753 aaaggatcttctgagatcccttttttctgcgcgtaaatctgctgcttgcaaaacaaaaaa 4812
Db 2645 AAAGGATCTTCTTGAGATCCTTTTCTCGCGGTAATCTGCTGCTTGCAACAACAAAAA 2586
QY 4813 ccaccgctaccagcgggtgtgttgcgggatacaagagctaccaactcttttccgaag 4872
|||||

Db 2585 CCACCGCTACCAGCGGTGGTTTGTTCGCCGATCAAGAGCTACCAACTCTTTTCCGAAG 2526
QY 4873 gtaactggcttcagcagagcgagataaccaataactgtccttctagtgtagccgtagtta 4932
Db 2525 GTAACGTGGCTTACGACAGCGCAGATACCAATACTGTCTTCTAGTGTAGCCGTAGTTA 2466
QY 4933 ggccaccacttcaagaactctgtagcaccgcctacatacctcgtctctgctaactcctgta 4992
Db 2465 GGCCACCACCTTCAAGAACTCTGTAGCACCCGCCCTACATACCTCGCTCTGCTAATCTGTTA 2406
QY 4993 ccagtggtcgtgccagtggcgataaagtctgtcttaccgggttgactcaagacgatatag 5052
Db 2405 CCAGTGGCTGCTGCCAGTGGCGATAAAGTCTGTCTTACCCGGTTGGACTCAAGACGATAG 2346
QY 5053 ttaccggataaaggcgagcgggtcgggctgaacgggggttcgtgcacacagccagccttg 5112
Db 2345 TTACCGGATAAAGCGCAGCGGTCCGGGTGAACGGGGGTTCGTGCACACAGCCACGCTTG 2286
QY 5113 gagcaacgacctaaccggaactgagatacctacagcgtgagctatgagaaaagcgccacg 5172
Db 2285 GAGCGAACGACCTACACCGAACTGAGATACCTACACGCTGAGCTATGAGAAAAGGCCACG 2226
QY 5173 cttcccgaaaggagaaaaagcggacaggtatcccgtaagcggcagggctcggaacagggagag 5232
Db 2225 CTTCCCGAAGGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGGGTCGGAACAGGAGAG 2166
QY 5233 cgcacgagggagcttccaggggggaaacgcctggtatatctttagtcccttgcgcacatg 5292
Db 2165 CGCACGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCTCTGTCGGGTTCGC 2106
QY 5293 cacctctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagccctatggaaa 5352
Db 2105 CACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATCGAAA 2046
QY 5353 aacgccagcaacgcggcctttttacgggttcctggccttttgtggccttttgcacatg 5412
Db 2045 AAGCCAGCAACCGCGCCTTTTACGGTTCTCTGGCCTTTTGTGGCCTTTTGTGTCACATG 1986
QY 5413 ttctttcctgcgttatccctgattctgtgtggataaacggtattaccgcctttgagtgagct 5472
Db 1985 TTCTTTCTCGCTTATCCCTTGATTCGTGGATAACCGTATTACCGCCTTTTGTGAGTGC 1926
QY 5473 gataccgctcgcgcagccgaacgacccgagcgagcgcagcagtcagtcagcgaggaagcgga 5532
Db 1925 GATACCGCTCGCCGACGCCGAACGACCCGAGCGCAGCGAGTCACTGAGCGAGGAACGGAA 1866
QY 5533 gagcgcctcaatacgcaaacccgctctctcccgcggttgccgattcattaatgacgctgg 5592
Db 1865 GAGCGCCCAATACGCCAAACCCGCTCTCCCCGCGGTGGCCGATTCAATTAATGCAGTGG 1806
QY 5593 cagcacaggtttcccgactggaaagcggcgagtcagtcagcgaacgcaattaatgtgagttag 5652
Db 1805 CACGACAGGTTTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAG 1746
QY 5653 ctcactcattaggcacccagcgttttacacttttatgcttccggctcgtatgttgtgtgga 5712
Db 1745 CTCACCTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGA 1686
QY 5713 attgtgagcggataaacaatttcacacagggaaaacagcgtatgacctgattacgcca 5767
Db 1685 ATTGTGAGCGGATACAATTTTCACACAGGAAACAGCTATGACCATGATTACGCCCA 1631

RESULT 5
US-08-990-442-1/c
; Sequence 1, Application US/089990442
; Patent No. 6090790
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Elof
; TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/990,442
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 310558.90028
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4283 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Plasmid DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: join(713..721, 981..1250)
US-08-990-442-1

Query Match 45.6%; Score 2631.8; DB 3; Length 4283;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3133 aattcaactggccgtcggttttacaaactgctgactgggaaacccctggcgttacccaactt 3192
|||||
Db 4265 AATTCACTGGCCGCTGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTT 4206

QY 3193 aatgccttgacgacatcccccttccgacagctggcgtaataagcgaagagccgcacc 3252
|||||
Db 4205 AATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGCCGCACC 4146

QY 3253 gatcgcccttcccaacagttgcgcagcctgaatggcgaaatggcgccctgatgcggtatttt 3312
|||||
Db 4145 GATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGSCGAATGGCCCTGATGCGGTATTTT 4086

QY 3313 ctcccttacgcactctgcggtattttcacacccgcataatggtgcactctcagtacaactctgc 3372
|||||
Db 4085 CTCCTTACGCATCTGTGCGGTATTTTCACACCCGCATATGGTGCACTCTCAGTACAATCTGC 4026

QY 3373 tctgatgcgcatagttaagccagccccgcacacccgcgaacaccccgctgacgcgcccctga 3432
|||||
Db 4025 TCTGATGCCGCATAGTTAAGCCAGCCCGACACCCGCCAACACCCCGCTGACGCCGCTGA 3966

QY 3433 cgggcttgctgctcccgccatccgcttacagacaaagctgtgaccgtctccgggagctgc 3492
|||||
Db 3965 CGGGCTTGCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC 3906

QY 3493 atgtgtcagaggttttcccgctcatcacccgaacgcgcgagacgaaggccctcgtgata 3552
|||||
Db 3905 ATGTGTCAGAGGTTTTCACCGTCATCACCGAAGCGCGGAGACGMAAGGGCCCTCGTGATA 3846

QY 3553 cgcctattttataggttaatgtcatgataataatggtttcttagacgtcaggtggcact 3612
|||||
Db 3845 CGCCTATTTTATAGGTTAATGTTCATGATAATAATGGTTTCTTAGACGTGAGTGGCACT 3786

QY 3613 ttctcggggaaatgtgcgcggaaacccctattttgtttatttttctaaatacatattcaaatatg 3672
|||||

Db 3785 TTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTCAAAATATG 3726

QY 3673 tatccgctcatgagacaataaaccttgataaaatgcttcaataatattgaaaaaaggaagagt 3732
|||||
Db 3725 TATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAT 3666

QY 3733 atgagtattcaacatttcggtgctgcccccttatcccttttttgcggcatttttgccttccct 3792
|||||
Db 3665 ATGAGTATTCAACATTTCCGTGTGCCCTTATTCCCTTTTTCGGGCATTTTGCCTTCCT 3606

QY 3793 gttttgtcacccagaaacgctggtgaaagtataaagtgcctgaagatcagttggtgca 3852
|||||
Db 3605 GTTTTGTCTACCCAGAAACGCTGGTGAAGTAAGATGCTGAAGATCAGTTGGGTGCA 3546

QY 3853 cgagtgggttacatcgaaactggatctcaacagcggtaagatcccttgagagtttttgcctccc 3912
|||||
Db 3545 CGAGTGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC 3486

QY 3913 gaagaacgttttccaatgatgagcaccttttaaagttctgctatgtggcgcggtattatcc 3972
|||||
Db 3485 GAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATATCC 3426

QY 3973 cgtattgacgcgcgggcaagagcaactcggcgccgcatacacactattctcagaatgacttg 4032
|||||
Db 3425 CGTATTGACGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTG 3366

QY 4033 gttgagtactcacaccagtcacagaaaaagcatctttacggatggcatgacagtaagaatta 4092
|||||
Db 3365 GTTGAGTACTCACCACTACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA 3306

QY 4093 tgcagtgctgccataaacatgagtataaacactcggcgaacttacttctgacaacgac 4152
|||||
Db 3305 TGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCAACTTACTTCTGACAAACGATC 3246

QY 4153 ggaggaccgaaggagcctaaccgcttttttgcacaaacatggggatcatgttaactcgcctt 4212
|||||
Db 3245 GGAGGACCGAAGGAGCTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCCTT 3186

QY 4213 gatcgttgggaaccggagctgaatgaagccataaccacacgacgagcgtgacacacgagtg 4272
|||||
Db 3185 GATCGTTGGGAACCCGGAGCTGAATGAAGCCATACCAACACGACGAGCGTGACACCACGATG 3126

QY 4273 cctgtagcaatggcaacaacgcttgcgcaaaactatttaactggcgaactacttactctagct 4332
|||||
Db 3125 CCTGTAGCAATGCCAACAAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCT 3066

QY 4333 tcccggaacaataatagactggatggagcggaataaagtgcaggaccacttctgcgc 4392
|||||
Db 3065 TCCCGGCAACAATTAATAGACTGGATGGAGGGGGATAAAGTTGCAGGACCACCTTCTGCGC 3006

QY 4393 tcggcccttcgggctgctggtttattgctgataaaatcctggtaagccctcccgtatcgttagttatctac 4452
|||||
Db 3005 TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCT 2946

QY 4453 cgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgttagttatctac 4512
|||||
Db 2945 CGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTAC 2886

QY 4513 acgacggggagtcaggcaactatggatgaacgaataagacagatcgcctgagataggtgccc 4572
|||||
Db 2885 ACGACGGGGAGTCAGGCAACTATGGATGAACGAATAGACAGATCGCTGAGATAGGTGCC 2826

QY 4573 tcaactgattaagcatttgaactgtcagaccaagtttactcaatataatacttttagattgat 4632
|||||
Db 2825 TCACTGATTAAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT 2766

QY 4633 ttaaaacttcatttttaatttaaaaggatctaggtaagatccctttttgataatctcatg 4692
|||||
Db 2765 TTAAAACTTCAATTTTAAATTTAAAAAGGATCTAGTGAAGATCCTTTTGTGATAATCTCATG 2706

QY 4693 accaaaaatcccttaacgtgagtttctcgtccactgagcgtcagacccccgtagaaaagatc 4752
|||||
Db 2705 ACCAAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATC 2646

QY 4753 aaagatatcttcttgagatccctttttttctgctgtaaatctgctgcttgcacaaacaaaaa 4812
|||||
Db 2645 AAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTGCAACAAAAA 2586
QY 4813 ccacgcgtaccagcgggtgtttgtttgcgggatacaagagctaccactcttttccgaag 4872
|||||
Db 2585 CCACCGCTACCAGCGGTGGTTGTGTTGCGCGGATCAAGAGCTACCAACTCTTTTCCGAAG 2526
QY 4873 gtaactggcttcagcagagcgagataccaaataactgtccttctagtgtagccgtagta 4932
|||||
Db 2525 GTAACGTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCCGTAGTTA 2466
QY 4933 ggcacacacttcaagaactctgttagcaccgcctacatacctcgtctgtaacctctgta 4992
|||||
Db 2465 GGCACCACTTCAAGAACTCTGTAGACCGCCTACATACCTCGCTCTGCTAATCCTGTTA 2406
QY 4993 ccagtggctgctgccagtgccagataagtcgtgtctttaccgggttggactcaagacgatag 5052
|||||
Db 2405 CCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAG 2346
QY 5053 ttaccggataaaggcgagcgggtcggggtgaacgggggggttcgtgcacacagccagcttg 5112
|||||
Db 2345 TTACCGGATAAGGCGCAGCGGTTCGGGTGAACGGGGGTTCTGTGCACACAGCCAGCTTG 2286
QY 5113 gagcgacgacctacaccgaactgagatacctacagctgagctagctatgagaaagcgccacg 5172
|||||
Db 2285 GAGCGAAGGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCCACG 2226
QY 5173 ctcccgaaaggagaaaggcggaaggtatccggttaagcggcgaggggtcggaacaggagag 5232
|||||
Db 2225 CTTCCGAAGGGAGAAAGGCGACAGGTATCCGGTAAGCGGCGAGGTCGGAACAGGAGAG 2166
QY 5233 cgcacgaggagcttccagggggaacgcctgggtatctttatagtcctgtcgggttctgc 5292
|||||
Db 2165 CGCAGGAGGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTCTGTCGGGTTTCG 2106
QY 5293 cacctctgacttgagcgtcgatctttgtgtagctcgtcagggggcgagcctatggaaa 5352
|||||
Db 2105 CACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAAGGGGGGCGGAGCCTATCGAAA 2046
QY 5353 aacgcaggaacgcggcctttttacgggttcctggccttttgcggccttttgcacatg 5412
|||||
Db 2045 AACGCCAGCAACGCGGCCTTTTACGGTTCCTGGCCTTTTGTGCTGGCCTTTTGTCTACATG 1986
QY 5413 ttcttctcgtgttatccctgattctgtggataaccgtattaccgcctttgagtgagct 5472
Db 1985 TTCTTTCTCGCTTATCCCTGATTTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCT 1926
QY 5473 gatacgcctcgcgcagccgaacgagcagcagcagcagcagcagcagcagcagcagcagc 5532
|||||
Db 1925 GATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAACGGGAA 1866
QY 5533 gagcgcccaatacgcgaacccgcctctcccccgcgcttgcccgagcagcagcagcagcag 5592
|||||
Db 1865 GAGCGCCCAATACGCAACCGCCTCTCCCGCGCTTGGCCGATTCATTAATGCAGCTGG 1806
QY 5593 cagcaaggtttcccgactggaagcgggcagtgagcgcgaacgcgaatgaatgtgagttag 5652
|||||
Db 1805 CACGACAGGTTTCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAG 1746
QY 5653 ctactcaattaggcacccaggctttacactttatgcttccggctcgtatgtgtgga 5712
Db 1745 CTCACATATTAGGCACCCCGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGA 1686
QY 5713 attgtgagcgggataacaattttcacacagggaacagctatgacctgattacgcca 5767
Db 1685 ATTGTGAGGGGATAACAATTTACACAGAGGAACAGCTATGACCATGATTACGCCA 1631

RESULT 6

US-08-659-206A-4/c

; Sequence 4, Application US/08659206A

; Patent No. 5922685
; GENERAL INFORMATION:
; APPLICANT: Rakmilevich, Alexander
; TITLE OF INVENTION: IL-12 Gene Therapy of Tumors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,206A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 110229.91144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6295 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "plasmid pWRG3196"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(955..1260, 1334..1675)
; OTHER INFORMATION: /product= "p35 gene product"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2377..3384
; OTHER INFORMATION: /product= "p40 gene product"
; US-08-659-206A-4

Query Match 45.6%; Score 2630; DB 2; Length 6295;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3138 actggcgcgtcgcttttacaaacgctcgtgactggggaacccctggcgttacccaacttaacg 3197
|||||
Db 6295 ACTGGCCGTCGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATCG 6236
QY 3198 ccttgacgacatccccctttcgcagctggtgtaataagcgaagagccgcaccgatcg 3257
|||||
Db 6235 CCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCG 6176
QY 3258 cccttcccaacagttgcgcagcctgaatggggaatggcctgatgcggtattttctcct 3317
|||||
Db 6175 CCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTGATGCGGTATTTCTCCT 6116
QY 3318 tacgcactctgctgggtattttcacacccgcataatgggcactctcagtcacataatctgctcga 3377
|||||
Db 6115 TACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACTCTCAGTACAACTTGCTCTGA 6056
QY 3378 tgccgcgatagttaagccagccccgcacacccgcgaacacccgcgtgacgcgcctgacgggc 3437
|||||
Db 6055 TGCCGCATAGTTAAGCCAGCCCCCGACACCCCGCTGACGGCCCTGACGGGC 5996
QY 3438 ttgtctgctcccgccatcccgcttacagacaagctgtgaccgtctccgggagctgcatgtg 3497
|||||

Db 5995 TTGTCGTCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTG 5936

QY 3498 tcagaggtttttcacagtcacatcaccgaaacgcgagagacgaaaaggccctcgtgatacgcct 3557

Db 5935 TCAGAGGTTTTCACCGTCATCACCGAAACCGCGGAGACGAAAGGCCCTCGTGATACGCCCT 5876

QY 3558 atttttataggttaatgctcatgataataaatgggtttcttagacgctcaggtggaacttttcg 3617

Db 5875 ATTTTATAGGTTAATGTGCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCG 5816

QY 3618 gggaaatgtcgcggaacccctatttgtttatttttctaaatatattgaaatacattcaaatatgtatcc 3677

Db 5815 GGGAAATGTGCGGGAACCCCTATTGTGTTATTTTCTAAATACATTCAATATATGTATCC 5756

QY 3678 gctcatgagacaataaacctgataaaatgcttcaataataattgaaaaaggaagagtatgag 3737

Db 5755 GCTCATGAGACAATAAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAG 5696

QY 3738 tattcaacatttcctgctgcgcccttattcccttttttgcggcattttgccttcctgtttt 3797

Db 5695 TATTCAACATTTCCGTGTGCCCCCTTATTCCCTTTTGTGCGGCATTTTGGCCTTCTGTTT 5636

QY 3798 tgctcaccagaaaacgctggtgaaagtataaagtgcctgaagatcagttgggtgcacgagt 3857

Db 5635 TGCTCACCCAGAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGT 5576

QY 3858 gggttacatcgaaactggatctcacaacagcggtgaagatccctgagagttttcgccccgaaga 3917

Db 5575 GGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCCCCGAAGA 5516

QY 3918 acgttttccaatgatgagcaccttttaaagttctgctatgtggcgcggtattatcccgatat 3977

Db 5515 ACGTTTTCCAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGCGGTTATTATCCCGTAT 5456

QY 3978 tgacgcggggcaagagcaactcggctgcgcgcgcatacacactattctcagaatgacttggtga 4037

Db 5455 TGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGA 5396

QY 4038 gtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaattatgcag 4097

Db 5395 GTACTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATGCAG 5336

QY 4098 tgctgccataaacctagagtataaacactcggcccaacttacttctgacaaacgctcggagg 4157

Db 5335 TGCTGCCATAAACCATGAGTGATAACACTCGCGGCCAACTTACTTCTGACAAACGATCGGAGG 5276

QY 4158 accgaaggagctaaccgcttttttgcacaacatggggggatcatgtaaactcgccttgatcg 4217

Db 5275 ACCGAAGGAGCTAACCGCTTTTGTGCACAACATGGGGGATCATGTAACTCGCCTTGATCG 5216

QY 4218 ttgggaacgggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgt 4277

Db 5215 TTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACGATGCCTGT 5156

QY 4278 agcaatggcaacaacgctgcgcgaaactattaactggcgaaactacttactctagcttcccg 4337

Db 5155 AGCAATGGCAACAACTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCG 5096

QY 4338 gcaacaattaatagactggatggaggcgggataaaagtbgcaggaccacttctgcgctcggc 4397

Db 5095 GCAACAATTAAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACITTCCTGCGCTCGGC 5036

QY 4398 ccttcgggctggctgggtttattgctgataaaatcgtgagccggtgagcgtgggtctcgcgg 4457

Db 5035 CCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGG 4976

QY 4458 tatcattgcagcactcgggcccagatggtaagccctcccgtatcgtagttatatcacgac 4517

Db 4975 TATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAC 4916

QY 4518 ggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagtgccctcact 4577

Db 4915 GGGGAGTCAGGCCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACT 4856

QY 4578 gattaagcattggttaactgtcagaccaagtttactactatatactatttagattgtataa 4637

Db 4855 GATTAAAGCATTGGTAACCTGTCTCAGACCAAGTTTACTCATATATACTTTAGATTGATTAAA 4796

QY 4638 acttcatttttaattaaaaagatctaggtgaagatccttttttgataatctcatgaccaa 4697

Db 4795 ACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCATGACCNA 4736

QY 4698 aatcccttaacgtgagtttctcgttccactgagcgtcagaccccgtagaaaaagatcaaaag 4757

Db 4735 AATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGG 4676

QY 4758 atcttcttgagatcctttttctgcgcgtaaatctgctgcttgcaaaacaaaaaccacc 4817

Db 4675 ATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAACCACC 4616

QY 4818 gctaccagcggtggtttgtttgcccggatcaagagctacaaactcttttccgaaggtaac 4877

Db 4615 GCTACCAGCGGTGTTGTTGTCGGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAAC 4556

QY 4878 tggcttcagcagagcgagatataccaaatactgtccttctagtgtagccgtagtgagcca 4937

Db 4555 TGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTCTAGTGTAGCCGTAGTAGGCCA 4496

QY 4938 ccaacttcaagaactctgtagcaccgcctacatacctcgtcctgctgtaactcctgttaccagt 4997

Db 4495 CCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGT 4436

QY 4998 ggtgctgcccagtggcgataaagtcgtgtcttaccgggttgagactcaagacgataagttacc 5057

Db 4435 GGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACC GGGTGGACTCAAGACGATAGTTACC 4376

QY 5058 ggataaaggcgcagcgttcgggctgaacgggggttcgtgcaacagcccagcttggagcg 5117

Db 4375 GGATAAGGCGCAGCGCTCGGGCTGAACGGGGGTTCTGTCACACACAGCCCAGCTTGGAGCG 4316

QY 5118 aacgacctacaccgaactgagatagatacctacagcgtgagctatgaaaaagcgcacgttcc 5177

Db 4315 AACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCGCCACGCTTCC 4256

QY 5178 cgaaggagaaaaggcgacaggtatccggttaaggcgaggggtcggaacaggagagcgcac 5237

Db 4255 CGAAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGGGTTCGGAACAGGAGAGCGCAC 4196

QY 5238 gagggagcttccagggggaaaacgcctgggtatctttatagtcctgtcgggttttcgcacct 5297

Db 4195 GAGGGAGCTTCCAGSGGAAACGCTGTTATCTTTTATAGTCTCTGTCGGGTTCGSCACCT 4136

QY 5298 ctgacttgagcgtcgatttttgatgctcgtcagggggcgagcctatggaaaaacgc 5357

Db 4135 CTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGSGGGGGGAGCCCTATGGAAAAACGC 4076

QY 5358 cagcaacgcggcctttttacgggttccctggccttttctggccttttctgcacatgcttctt 5417

Db 4075 CAGCAACGCGGCCCTTTTACGGTTCTCTGGCCCTTTTGTGCTGGCCCTTTTGTCTACATGTTCTT 4016

QY 5418 tctgctgttatccccctgattctgttgataaacctgtattaccgccttttgagtgagctgatac 5477

Db 4015 TCCTGCGTTATCCCTCGATTCTGTGGATAACCCGTATTACCGCCTTTGAGTGAGCTGATAC 3956

QY 5478 cgctcgcgcgcagccgaacgacgcgagcgcagcgtcagtgagcgaaggcgaagagcg 5537

Db 3955 CGCTCGCGCAGCCGAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAAGCGGAAGAGCG 3896

QY 5538 cccaatacgcaaacccctctccccgcgcgcttgccgattcattaatgcagctggcacga 5597

Db 3895 CCCAATACGCAAAACCCGCTCTCCCCCGCGCTTGGCCGATTCTAATATGCAGCTGGCACGA 3836

QY 5598 cagggtttcccgactggaaaagcgggagtgagcgaacgcgaattaatgtgagttagctcac 5657

Db 3835 CAGGTTTCCCGACTGGAAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCAC 3776

QY 4335 ccggcaacaataatagactggatggaggcgagataaaagtgcaggaccacttctgcgctc 4394
|||||
Db 1201 CCGGCAACAATTAAAGACTGGATGGAGGCGGATAAAGTTGCAGGACCATTCTGCGCTC 1260
QY 4395 ggccttcgcgctggtgtttattgctgataaaatctggagccggtgagcgtgggtctcg 4454
|||||
Db 1261 GGCCCTTCGGCTGGCTGTTATTGCTGATAAATCTGGAGCCGTTGAGCGTGGTCTCG 1320
QY 4455 cggatatcattgcagcactggggccagatggtaagccctcccgtatcgttagttatctacac 4514
|||||
Db 1321 CGGTATCATTTGCAGACTGGGGCCAGATGGTAAGCCCTCCGTTATCGTAGTTATCTACAC 1380
QY 4515 gacggggagtcaggcaactatggtgaacgaataagacagatcgctgagataggtgcctc 4574
|||||
Db 1381 GACGGGGAGTCAGGCAACTATGATGAACGAATAGACAGATCGCTGAGTAGGTGCCTC 1440
QY 4575 actgattaagcattggttaactgtcagaccgaagtttactcatatatactttagattgatt 4634
|||||
Db 1441 ACTGATTAAGCATTTGGTAACGTGTGAGACCCNAAGTTACTCATATATACTTTAGATTGATT 1500
QY 4635 aaaacttcatttttaatttaaaggatctaggtgaagatcccttttttgataaatctcatgac 4694
|||||
Db 1501 AAAACTTCATTTTAAATTTAAAGGATCTAGTGAAGATCCCTTTTGTGATAATCTCATGAC 1560
QY 4695 caaaatcccttaacgtagtgttttcgttccactgagcgtcagacccccgtagaaaaagatcaa 4754
|||||
Db 1561 CAAATCCCTTAAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGATCAA 1620
QY 4755 aggatcttcttgagatccctttttctgcggttaactgtcgtgttgcaacaaaaaac 4814
|||||
Db 1621 AGGATCTTCTTGAGATCCCTTTTCTGCGGTAATCTGCTGTGCAACAAAAAACC 1680
QY 4815 accgtaccagcgtggtttgtttgccggatcaagagctaccactcttttccgaaggt 4874
|||||
Db 1681 ACCGCTACCAGCGGTGTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGT 1740
QY 4875 aactggcttcagcagagcgcagataaccaataactgtccttctagtagcgttagttagg 4934
|||||
Db 1741 AACTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTTCTTCTAGTGTAGCCGTAGTAGG 1800
QY 4935 ccaccacttcagaactctgtagcaccgcctacatacctcgtcgtgtaaatcctgttacc 4994
|||||
Db 1801 CCACCACCTTCRAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCTCTTACC 1860
QY 4995 agtggctgctgccagtgccgataaagtctgtcttaccgggttgactcaagacgatatagt 5054
|||||
Db 1861 AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGTTGGACTCAAGACGATAGTT 1920
QY 5055 accggataaggcgagcgtcggtggaacgggggttcgtgcacacagccagccttgga 5114
|||||
Db 1921 ACCGGATAAGGCGAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAGCCACGCTTGG 1980
QY 5115 gcgaacgacctacacccgaactgagataacctacagcgtgagctatgagaagcgccacgct 5174
|||||
Db 1981 GCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCT 2040
QY 5175 tcccgaagggaaggaaggcggacaggtatccgggtaagcggcagggttcggaacaggagcgc 5234
|||||
Db 2041 TCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCG 2100
QY 5235 cacgagggagcttccaggggggaacgccttggtatctttatagtcctgctggggtttcgcca 5294
|||||
Db 2101 CACGAGGGAGCTTCCAGGGGGAAACGCCCTTGGTATCTTTATAGTCTGTGCGGTTTCGCC 2160
QY 5295 cctctgacttgagcgtcgattttttgtgtagtctgtcgtcagggggcgagcctatggaaaaa 5354
|||||
Db 2161 CCTCTGACTTGAGCGCTCGATTTTGTGTATGCTCGTCAAGGGGGCGGAGCCTATGGAAAAA 2220
QY 5355 cgccagcaacgggccttttttacggttctcgtgaccttttgcggccttttgcctcacatgtt 5414
|||||
Db 2221 CGCCAGCAACGGGCCCTTTTACGGTTCTTGGCCCTTTTGTGTCGCTTTTGCTCACATGTT 2280
QY 5415 ctttcctgcgttatccctctgattctgtggataaacgctgattaccgccttttgagtgcgctga 5474

Db 2281 CTTTCTCGGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGA 2340
QY 5475 taccgctcgcgcagcgcgaacgaccgagcgcgagcagtcagtgagcgcgagcgcggaaga 5534
|||||
Db 2341 TACCGCTCGCCGACGCGAAACGACCGAGCGCAGCAGTCAGTGAGCGAGGAACGGAAGA 2400
QY 5535 gcgccaatacgcgaacccgcctctccccgcgcgttggccgattcatttaataatgcagctggca 5594
|||||
Db 2401 GCGCCCAATACGCAAAACCGCCTCTCCCCCGCGTGGCCGATTCAATTAATGAGCTGGCA 2460
QY 5595 cgacaggtttcccgactggaaagcgggcagtcagtcagcgcgaacgcaataatagtgcagttagct 5654
|||||
Db 2461 CGACAGGTTTCCCGACTGGAAGCGGGCAGTCAGCGCAACGCAATTAATGTGAGTTAGCT 2520
QY 5655 cactcattaggcacccaggcttttacactttatgcttccggctcgtatgtgtgtggaat 5714
|||||
Db 2521 CACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGAAT 2580
QY 5715 tgtgagcggataacaatttcacacaggaacacgctatgaccatgattacgcca 5767
|||||
Db 2581 TGTGAGCGGATAACAATTTACACAGGAACAGCTATGACCATGATTACGCCA 2633

RESULT 8
US-08-789-333F-58/C
; Sequence 58, Application US/08789333F
; Patent No. 6153380
; GENERAL INFORMATION:
; APPLICANT: No. 6153380an, Garry P
; APPLICANT: Rothenberg, S. M.
; TITLE OF INVENTION: METHODS FOR SCREENING FOR TRANSDOMINANT INTRACELLULAR
; TITLE OF INVENTION: EFFECTOR PEPTIDES AND RNA MOLECULES
; FILE REFERENCE: A642601DJBRMSDSS
; CURRENT APPLICATION NUMBER: US/08/789,333F
; CURRENT FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: 08/589,108
; PRIOR FILING DATE: 1996-01-23
; PRIOR APPLICATION NUMBER: 08/589,911
; PRIOR FILING DATE: 1996-01-23
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 4950
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: retroviral
; OTHER INFORMATION: vector with presentation construct sequence.
US-08-789-333F-58

Query Match 45.6%; Score 2628; DB 3; Length 4950;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3137 cactggcgcgtcgttttacaaacgctcgtgactgggaaaaacccctggcgttacccaacttaatc 3196
|||||
Db 4675 CACTGGCCGCTCGTTTACAAACGTCGTGACTGGGAAAAACCCCTGGCGTTACCCAACTTAATC 4616
QY 3197 gccttgacacacatccccctttcgcagcctgaatggcgaaatggcgctgatgcggtattttctcc 3256
|||||
Db 4615 GCCTTGACGACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAGAGGCCGCCACCGATC 4556
QY 3257 gcccttcccaaacagttgcgcagcctgaatggcgaaatggcgctgatgcggtattttctcc 3316
|||||
Db 4555 GCCCTTCCCAACAGTTGCGCAGCCTGAATGGGAATGGCGCTGATGCGGTATTTCTCC 4496
QY 3317 ttacgcattctgctgggtatttcacacccgcataatggcgactctcagtacaatctgctctg 3376
|||||
Db 4495 TTACGCATCTGTGCGGTATTTCACACCCGCATATGGTGCACTCTCAGTACAATCTGCTCTG 4436
QY 3377 atgccgcgcatagtttaagccagcccccgacaccccgcaacccccgctgacgcgcctgacggg 3436

||||| 4435 ATCCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCCGCTGACGGCCCTGACGGG 4376

QY 3437 cttgtctcgtcccgcatccgcttacagacaagctgtgaccgtctccggagctgcatgt 3496

Db 4375 CTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCTGCATGT 4316

QY 3497 gtcagagggttttcaccgctcatcaccgaaacgcgcgagacgaaaggcctcgtgatacgc 3556

Db 4315 GTCAGAGGTTTTACCCGTCATCACCGAAACGCGGAGACGAAAGGGCCTCGTGATACGCC 4256

QY 3557 tattttataggttaatgtcatgataataatggttctcttagacgtcaggtggcacttttc 3616

Db 4255 TATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTC 4196

QY 3617 ggggaaatgtgcgcggaacccctatttggtttatttttctaaatacatattcaaatatgtatc 3676

Db 4195 GGGGAAATGTGCGCGGAACCCCTATTTCGTTTATTTCTTAAATACATTCAAATATGTATC 4136

QY 3677 cgctcatgagacaataaacccctgataaaatgcttcaataatattgaaaaagggaagagtatga 3736

Db 4135 CGCTCATGACACAATAACCCCTGATAATGCTTCAATAATATTGAAAAAGGAAGATATGA 4076

QY 3737 gtattcaaatctccgtgtgcgccttattcccttttttgcggcattttgccttccctgttt 3796

Db 4075 GTATTCAACATTTCCGTGTGCGCCCTTATTTCCTTTTGTGCGGCATTTTGCCTTCTCTGTTT 4016

QY 3797 ttgctcacccagaaaacgctggtgaaagtataaagatgctgaaagatcagttgggtgcacgag 3856

Db 4015 TTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAG 3956

QY 3857 tgggttacatcgaaactggatctcaacgcgcggtgaagatcccttgagagttttcgccccgaag 3916

Db 3955 TGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAG 3896

QY 3917 aacgttttccaatgatgagcacacttttaaaagtctctgatatgtggcgcggtattatcccgta 3976

Db 3895 AACGTTTTCCAATGATGAGCACATTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTA 3836

QY 3977 ttgacgcgcgggcaagagcaactcggtcgcgcgcatacacactattctcagaaatgacttggttg 4036

Db 3835 TTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTTGGTTG 3776

QY 4037 agtactcacagtcacagaaaaagcatcttaCggatggcatgacagtaagagaattatgca 4096

Db 3775 AGTACTCACAGTCCACAGAAAGCATCTTACGGATGGCATGGCATGACAGTAAGAGAAATTATGCA 3716

QY 4097 gtctgccataaccatgagtgtataaactgcggccaaacttacttctgacaacgatcggag 4156

Db 3715 GTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAG 3656

QY 4157 gaccgaaggagctaaccgcttttttgcaacaacatgggggatcatgtaaactcgcccttgatc 4216

Db 3655 GACCGAAGGAGCTAACCGCTTTTTCACACAACATGGGGATCATGTAACTCGCCTTGATC 3596

QY 4217 gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctg 4276

Db 3595 GTTGGGAACCGGAGCTGAATGAAGCCCATACCAACGACGAGCGGTGACACCACGATGCCTG 3536

QY 4277 tagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctagcttccc 4336

Db 3535 TAGCAATGGCAACAACAGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC 3476

QY 4337 ggcaacaattaatagactggatggagcggaataaaagtgcaggaccacttctgcgctcgg 4396

Db 3475 GGCAACAATTAATAGACTGGATGGAGGCGGATAAAAGTTGCAGGACCCTTCTTGGCCTCGG 3416

QY 4397 cccttcgggctggcctggtttattgtctgataaaatctggagccggtgagcgtgggtctcgcg 4456

Db 3415 CCCTTCCGGCTGCGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGGTGGGTCTCGCG 3356

QY 4457 gtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacga 4516

|||||

Db 3355 GTATCATTTGCACACTGGGGCCAGATGGTAAGCCCTCCCCTATCGTAGTTATCTACACGA 3296

QY 4517 cggggagtcaggcaactatggatgaacgaataagacagatcgctgagataggtgcctcac 4576

Db 3295 CGGGGAGTCAGCAACTATGGATGAACGAATAGACAGATCGTGAGATAGGTGCCTCAC 3236

QY 4577 tgattaagcattggttaactgtcagaccaaagtttactcataataatacttttagattgatttaa 4636

Db 3235 TGATTAAGCATTTGGTAACCTGTCAGACCAAGTTTACTCATATATATCTTTTAGATTGATTAA 3176

QY 4637 aacttcatttttaatttaaaaggatcttagtgaagatcccttttgataatctcatgacca 4696

Db 3175 AACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCATGACCA 3116

QY 4697 aaatcccttaacgtgagttttcttcgcgcgtaaatctgctgtgcacacacacacacacac 4756

Db 3115 AAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGNAAGATCAAAAG 3056

QY 4757 gatctcttgagatccctttttcttcgcgcgtaaatctgctgtgcacacacacacacacacac 4816

Db 3055 GATCTTCTTGAGATCCTTTTTCGCGGTAATCTGCTGCTTGCAAAACAAAAAACCCAC 2996

QY 4817 cgctaccagcggtgggtttgtttgcgggatacaagagctaccaactcttttccgaaggtaa 4876

Db 2995 CGCTACCAGCGGTGGTTTGTTCGCGGATCAAGAGCTACCAACTCTTTTTCGAAAGGTAA 2936

QY 4877 ctggcttcagcagagcgcagataaccaataacttgccttcttagtgcgtagttaggcc 4936

Db 2935 CTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTCTTAGTGTAGCCGTAGTAGGCC 2876

QY 4937 accacttcaagaactctgtagcaccgcctacataacctgcctgcctgctaaatcctgttaccag 4996

Db 2875 ACCACTTCAAGAACTCTGTAGCACCCGCTACATACTCGCTCTGCTTAATCTCTGTACCAG 2816

QY 4997 tggctgctgccagtggcgataaagtctgttcttaccgggttggaactcaagacgatagttac 5056

Db 2815 TGGCTGCTGCCAGTGGCGATAAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAGTTAC 2756

QY 5057 cggataaggcgacgcggtcgggctgaacgggggttcgtgcacacagccccagcttggagc 5116

Db 2755 CGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGTTCTGTGCACACAGCCACGCTTGGAGC 2696

QY 5117 gaacgaacctacacccgaactgagataacctacagcgtgagctatgagaaaaagccacgccttc 5176

Db 2695 GAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAAGCCACGCTTC 2636

QY 5177 ccgaaggggagaaaaggcggaaggtatccggtaacggcgaggggtcggaaacagagagcgca 5236

Db 2635 CCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCA 2576

QY 5237 cgaggagcttccagggggaaacgcctggttatctttatagtcctgtcggggttcgccaacc 5296

Db 2575 CGAGGGAGCTTCCAGGGGGAAACGCTGTTATAGTCTCTGCTCGGGTTTCGCCACC 2516

QY 5297 tctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagccctatggaaaaacg 5356

Db 2515 TCTGACTTGAGCGTCGATTTTGTGATGCTCTCAGGGGGCGGAGCCTATGGAACAAACG 2456

QY 5357 ccagcaacgcggcctttttacggttccctggccttttgcctgttctgctcacatgttct 5416

Db 2455 CCAGCAACGCGGCCCTTTTACGGTTCTCTGGCCCTTTTGTGGCCTTTTGTCTCACATGTTCT 2396

QY 5417 ttccctggttatccctgattctctgttgataaacggtattaccgcctttgagtgagctgata 5476

Db 2395 TTCTTGGGTTATCCCTTGATTTCTGTGGATAACCGTATTACCGCCTTTTGTAGTGAGCTGATA 2336

QY 5477 ccgctcgcgcagccgaacgacgcgagcgcagcagtcagtcagtcaggaagcgaagagc 5536

Db 2335 CCGCTCGCCGACGCGAAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGGAGAGC 2276

QY 5537 gcccaatacgcaaacgcgcctctccccgcgcgttggccgattcattaatgcagctggcacg 5596

Db 2275 GCCCAATACGCCAAACCGCCTCTCCCCCGCGCTTGSCCGGATTCATTAAATGCAGCTGGCACG 2216

QY	5597	acagggttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctca	5656
Db	2215	ACAGGTTTCCCGACTGTAAGCGGGCAGTGAGCGCAACGAATTAATGTGAGTAGCTCA	2156
QY	5657	ctcattaggcaccacggcctttacacctttatgcttccggctcgtatgttggtggaattg	5716
Db	2155	CTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTG	2096
QY	5717	tgagcggataacaatttcacacaggaaacagctatgaccatgattacg	5764
Db	2095	TGAGCGGATAACAATTTACACAGGAAACAGCATGACCCATGATTACG	2048

```

RESULT          9
US-09-147-208-26/c
; Sequence 26, Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
;   APPLICANT:
;   TITLE OF INVENTION: Antiviral Ricin-Like Proteins
;   NUMBER OF SEQUENCES: 71
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: BERESKIN & PARR
;     STREET: 40 King Street West
;     CITY: Toronto
;     STATE: Ontario
;     COUNTRY: Canada
;     ZIP: M5H 3Y2
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/147,208
;     FILING DATE: 02-MAR-1999
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Rudolph, John R.
;     REGISTRATION NUMBER: 38,003
;     REFERENCE/DOCKET NUMBER: 7841-76
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (416) 364-7311
;       TELEFAX: (416) 361-1398
;     INFORMATION FOR SEQ ID NO: 26:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 9639 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
US-09-147-208-26

```

Query Match 45.6%; Score 2628; DB 4; Length 9639;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2628; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

Db	9457	TTACGCATCTGTGCGGTATTTACACACCGCATATGGTGCAC	TCTCAGTACAATCTGCTCTG	9399
QY	3377	atgcgcgatagttaagccagcccgacacccgcgaaccccgctgacgcgccttgacggg		3436
Db	9397	ATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGGCCCTGACGGG		9338
QY	3437	cttgctgctccccggcctacagacaagctgtgacgcgtctccgggagctgcgtatgt		3496
Db	9337	CTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGT		9278
QY	3497	gtcagaggttttcaccgctcatcccgaaacgcgcgagacgaagggcctcgtgatacgc		3556
Db	9277	GTcAGAGGTTTTCACCGTCAACCCGAAACGCGGAGACGAAAGGGCCTCGTGATACGCC		9218
QY	3557	tattttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttc		3616
Db	9217	TATTTTATAGGTTAATGTcATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTC		9158
QY	3617	ggggaaatgtgcgcggaaacccctattgtttatttttctaatacatctcaaatatgtatc		3676
Db	9157	GGGGAATGTGCGCGGAACCCCTATTGTGTTATTTTCTAAATACATTCAAAATATGATC		9098
QY	3677	cgctcatgagacaataaacctgataaatgcttcaataatatgaaaaaggaagagtatga		3736
Db	9097	CGTcATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGATATGA		9038
QY	3737	gtattcaacatttcgctgccccctattcccttttttgcgccattttgcccccttcctgttt		3796
Db	9037	GTATTCAACATTTCCGTGTGCCCCCTATTCCCTTTTTTTCGGGCATTTTGCCCTTCCGTGTT		8978
QY	3797	ttgctcaccgaaacgctggtgaaagtataaagatgctgaagatcagttgggtgcacgag		3856
Db	8977	TTGCTCACCCAGAAACGCTGGTGAAAGTAAAAAGATGCTGAAGATCAGTTGGGTGCACGAG		8918
QY	3857	tgggttacatcgaactggatcacaacagcggtaaagatccttgagagttttcgccccgaag		3916
Db	8917	TGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGTCTTGAGAGTTTTCGCCCCGAAG		8858
QY	3917	aacgttttccaatgatgagcaccttttaaagttctgctatgtggcgcggtattatccccga		3976
Db	8857	AACGTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTA		8798
QY	3977	ttgacgcggggcaagagcaactcggtcgccgcatacacatttctcagaatgacttgggtt		4036
Db	8797	TTGACGCCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGACTGGTTG		8738
QY	4037	agtaactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaaatgatga		4096
Db	8737	AGTACTCACCACTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAATTATGCA		8678
QY	4097	gtgctgccataaacatgagtgaataacactcgccgaacttacttctgacaaacgatcggag		4156
Db	8677	GTGCTGCCATAAACCATGAGTGATAAACACTCGGGCCAACCTACTTCTGACAAACGATCGGAG		8618
QY	4157	gaccgaaggagctaaccgcttttttgcaacaatgggggatcatgtaactcgcccttgatc		4216
Db	8617	GACCGAAGGAGCTAACCGCTTTTTCACACAACATGGGGGATCATGTAACTCGCCTTGATC		8558
QY	4217	gttgggaacggagctgaatgaagccataccaaacgacgagcgtgacacccacgatgcctg		4276
Db	8557	GTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACGATGCCCTG		8498
QY	4277	tagcaatggcaacaacgttgcgcaaaactattaaactggcgaactacttactctagcttccc		4336
Db	8497	TAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCC		8438
QY	4337	ggcaacaattaatagactggatggagcggaataaagttgcaggaccacttctgcgctcgg		4396
Db	8437	GGCAACAATTAATAGACTGGATGGAGCGGGATAAAGTTGCAGGACCACCTTCTGCGCTCGG		8378
QY	4397	cccttcggctggctggtttattgtgtataaaatctggagccggtgagcgtgggtctcgcg		4456
Db	8377	CCCTTCGGCTGGCTGGTTATTGCTGATAAAATCTGAGCCGGTGAGCGTGGGTCTCGCG		8318

QY 4684 aatctcatgacccaaaaatcccttaacgtgagtttctgtccactgagcgtcagaccccgta 4743
Db 5506 AATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTA 5447
QY 4744 gaaaagatcaaaaggatcttcttgagatcccttttttctgcgcgtaaatctgctgtgcaa 4803
Db 5446 GAAAAGATCAAAAGGATCTTCTTGAGATCCCTTTTCTGCGCGTAATCTGCTTGTCAA 5387
QY 4804 acaaaaaaacaccgctaccagcgggtgtttgtttgcccggatcaagagctaccaactctt 4863
Db 5386 ACAAAAAAACACCGCTACCGCGGTGTTTGTTCGGGATCAAGAGGTACCAACTCTT 5327
QY 4864 ttccgaaggtaaactggcttcagcagagcagagataccaaaataactgtccttctagtgtag 4923
Db 5326 TTTCCGAAGGTAACTGGCTTCAGCAGAGCCGAGATACCAATACTGTCTTCTAGTGTAG 5267
QY 4924 ccgtagttagccaccacttcaagaaactctgtagcaccgcctacatacctcgtctgcta 4983
Db 5266 CCGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTA 5207
QY 4984 atcctgttaccagtggtgctgctgccagtgccgataaagtcgtgtcttaccgggttgactca 5043
Db 5206 ATCCTGTTACCAAGTGGCTGCTGCCAGTGGGATAAGTCGTCTTACCGGTTGGACTCA 5147
QY 5044 agacgatgttaccggataaaggcgccagcggtcgggctgaacgggggttcgtgcacacag 5103
Db 5146 AGACGATAGTTACCGGATAAGCGCGACGCGTTCGGGCTGAACGGGGGTTCTGTGCACACAG 5087
QY 5104 cccagcttgagcgaacgacctacacccgaactgagatacctacagcgtgagctatgagaa 5163
Db 5086 CCCAGCTTGGAGCAACGACCTACACCCGAACCTGAGATACCTACAGCGTGAGCATTGAGAA 5027
QY 5164 agcgccacgcttcccgaaggaggagaaaaggcgacaggtatccggtaagcggcagggtcgga 5223
Db 5026 AGCGCCACGCTTCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCAGGGTCGGA 4967
QY 5224 acaggagagcagcaggagcttccagggggaacgcctggtatctttatagtctgtc 5283
Db 4966 ACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAAACGCCTGGTATCTTTATAGTCTCTGTC 4907
QY 5284 gggtttcgccacctgacttgactgagcgtcgatttttgtgatctcgtcagggggcgagc 5343
Db 4906 GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAAGGGGGCGGAGC 4847
QY 5344 ctatggaaaaacgccagcaacgcggcccttttacgggttcctggcccttttgctggccctttt 5403
Db 4846 CTATGGAAAAACGCCAGCAACGGGCCCTTTTACGGTTCTTGGCCCTTTTGTGGCCCTTTT 4787
QY 5404 gctcacatgttcttccctgcgttatccctcgtattctgtggataaacgtattaccgccttt 5463
Db 4786 GCTCACATGTTCTTCTCGGTTATCCCTCGANTCTGTGGATAACCGTATTACCGCCTTT 4727
QY 5464 gagtgagctgataccgctcgccgcagccgaacgcagcagcagcagcagcagcagcagcag 5523
Db 4726 GAGTGAGCTGATACCGCTCGCCGACGCGAAGCACCGGAGCGCAGCGAGTCAGTGAGCGAG 4667
QY 5524 gaagcgggaagcgcgcccaatacgcgaacccgctctcccgcggttgccgattcattaa 5583
Db 4666 GAAGCGGAAGAGCGCCCAATACGCAACCGCCTCTCCCGCGCGTTGGCCGATTCAATAA 4607
QY 5584 tgcagctggcacgacaggttcccgactggaagcgggcagtgagcgcgaacgcaattaat 5643
Db 4606 TGCAGCTGSCACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAAT 4547
QY 5644 gtgagttagctcactcattaggcaccgccgctttacactttatgcttcgggctcgtatg 5703
Db 4546 GTGAGTTAGCTCACTCATTAGGCACCCCGAGGCTTTACACTTTATGCTTCGGCTCGTATG 4487
QY 5704 ttgtgtggaattgtgagcgggataaacaatttccacaggaacagctatgaccatgattac 5763
Db 4486 TTGTGTGGAATTGTGAGCGGATAACAATTTTACACAGGAAACAGCTATGACCATGATTAC 4427
QY 5764 gcc 5766

Db 4426 GCC 4424
|||
RESULT 11
US-08-232-016-22/c
; Sequence 22, Application US/08232016
; Patent No. 5952547
; GENERAL INFORMATION:
; APPLICANT: CORNELISSEN, Marc
; APPLICANT: SOETAERT, Piet
; APPLICANT: STAM, Maïke
; APPLICANT: DOCKX, Jan
; APPLICANT: VAN AARSSEN, Roel
; TITLE OF INVENTION: MODIFIED GENES AND THEIR EXPRESSION IN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,016
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91402920.2
; FILING DATE: 30-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92400820.4
; FILING DATE: 25-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K
; REGISTRATION NUMBER: P39,300
; REFERENCE/DOCKET NUMBER: 010830-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: plasmid DNA designated as pJD884
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1869
; OTHER INFORMATION: /note= "Coding region of a
; OTHER INFORMATION: truncated bt2 (cryIAb) gene, also designated as the bt884
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1877..2110
; OTHER INFORMATION: /note= "3' regulatory sequence
; OTHER INFORMATION: containing the polyadenylation site derived from Agrobacte
; OTHER INFORMATION: T-DNA gene 7."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2480..3005
; OTHER INFORMATION: /note= "35S promoter sequence
; OTHER INFORMATION: derived from Cauliflower mosaic virus."
; FEATURE:
; NAME/KEY: misc_feature

Qy	4924	ccgtagtagccaccacattcaagaactctgtagcaccgacctacatacactcgtctgcta	4988
Db	5350	ccgtagtagccaccacattcaagaactctgtagcaccgacctacatacactcgtctgcta	5291
Qy	4984	atcctgttaccagtggtgctgctgccagtggcgataaagtcgtgtcttaccgggttggaactca	5043
Db	5290	atcctgtttaccagtggtgctgctgccagtggcgataaagtcgtgtcttaccgggttggaactca	5231
Qy	5044	agacgatatgttaccgggataaaggcgacggctggggctgaacgggggttcgtgcacacag	5103
Db	5230	agacgatatgttaccgggataaaggcgacggctggggctgaacgggggttcgtgcacacag	5171
Qy	5104	cccagcttggaggaacgacctacacggaactgagatacctacagcgtgagctatgagaa	5163
Db	5170	cccagcttggaggaacgacctacacggaactgagatacctacagcgtgagcattgagaa	5111
Qy	5164	agcgcacgcttcocgaaggagaaaaggcgacaggtatccggtaagcggcagggtcgga	5223
Db	5110	agcgcacgcttcccgaaaggagaaaaggcgacaggtatccggtaagcgcagggctcga	5051
Qy	5224	acaggagagcgacagaggagcttccagggggaaaaagcctggatatctttatagtcctgctc	5283
Db	5050	acaggagagcgacagaggagcttccagggggaaaaagcctggatatctttatagtcctgctc	4991
Qy	5284	gggtttcgcacacctctgacttgagcgtcgatttttgcgtgatgctcgtcagggggcggaagc	5343
Db	4990	gggtttcgcacacctctgacttgagcgtcgatttttgcgtgatgctcgtcagggggcggaagc	4931
Qy	5344	ctatggaaaaacgcccagcaacgcggcctttttacgggttcctggccttttgcctggcctttt	5403
Db	4930	ctatggaaaaacgcccagcaacgcggcctttttacgggttcctggccttttgcctggcctttt	4871
Qy	5404	gctcacatgttcttctctgcgttatccccctgatctctgtggataacggtattaccgccttt	5463
Db	4870	gctcacatgttcttctctgcgttatccccctgatctctgtggataacggtattaccgccttt	4811
Qy	5464	gagtgagctgataccgctcgccgcagccgaacgacccgagcgcagcagtgatcagtcgagcgag	5523
Db	4810	gagtgagctgataccgctcgccgcagccgaacgacccgagcgcagcagtgatcagtcgagcgag	4751
Qy	5524	gaagcggaaagagcggcccaatacgcgaacccgctctccccgcgcgttggccgatttcattaa	5583
Db	4750	gaagcggaaagagcggcccaatacgcgaacccgctctccccgcgcgttggccgatttcattaa	4691
Qy	5584	tgcagctggcacgacaggtttcccgactggaaaagcgggcagtgagcgcgaacgaattaat	5643
Db	4690	tgacagctggcacgacaggtttcccgactggaaaagcgggcagtgagcgcgaacgaattaat	4631
Qy	5644	gtgagttagctcactcattaggaaccccgaggtttacacatttatgcttcgggctcgtatg	5703
Db	4630	gtgagttagctcactcattaggaaccccgaggtttacacatttatgcttcgggctcgtatg	4571
Qy	5704	ttgtgtggaattgtgagcgggataacaattttcacaggaacagcgtatgaccaatgattac	5763
Db	4570	ttgtgtggaattgtgagcgggataacaattttcacaggaacagcgtatgaccaatgattac	4511
Qy	5764	gcc	5766
Db	4510	gcc	4508

```

; TITLE OF INVENTION: and Methods for Activating CD4+ T Cells
; FILE REFERENCE: TSRI 536.1
; CURRENT APPLICATION NUMBER: US/09/194,285
; CURRENT FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: PCT/US97/08697
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 60/018,175
; PRIOR FILING DATE: 1996-05-23
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4713
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-194-285-7

```

Query Match 45.5%; Score 2624.8; DB 4; Length 4713;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 2626; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	3137	cactggcgtcgcttttacaacgctgtagctgggaaaaacccctggcgttacccaacttaatc	3196
Db	2086	cactggcgtcgcttttacaacgctgtagctgggaaaaacccctggcgttacccaacttaatc	2145
QY	3197	gccttgccagcacatccccctttcgccagctggcgtaatagcgaagaggcccgccagcaccgac	3256
Db	2146	gccttgccagcacatccccctttcgccagctggcgtaatagcgaagaggcccgccagcaccgac	2205
QY	3257	gcccttcccaacagttgcgcagccctgaatggcggaatggcgccctgatgcgggtattttctcc	3316
Db	2206	gcccttcccaacagttgcgcagccctgaatggcggaatggcgccctgatgcgggtattttctcc	2265
QY	3317	ttacgcatactgtgcgggtattttcacacccgcataatatggtgcactctcagtaacaatctgctctg	3376
Db	2266	ttacgcatactgtgcgggtattttcacacccgcataatatggtgcactctcagtaacaatctgctctg	2325
QY	3377	atgccgcatagtttaagccagccccgcacacccgcgcacccccctgacgcgcacctgacggg	3436
Db	2326	atgccgcatagtttaagccagccccgcacacccgcgcacccccctgacgcgcacctgacggg	2385
QY	3437	cttgctgctccggcgcataccgcttacagacaagctgtgaccgtctctccgggagctgcatgt	3496
Db	2386	cttgctgctccggcgcataccgcttacagacaagctgtgaccgtctctccgggagctgcatgt	2445
QY	3497	gtcagagggttttcaccgctcatcacccgaaacgcgcgcagacgaaaaggccctcgtgatacgc	3556
Db	2446	gtcagagggttttcaccgctcatcacccgaaacgcgcgcagacgaaaaggccctcgtgatacgc	2505
QY	3557	tatttttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcacttttc	3616
Db	2506	tatttttataggttaatgtcatgataataatgggtttcttagacgtcaggtggcacttttc	2565
QY	3617	ggggaaaatgtgcgcggaaacccctatttgtttatttttctaaatacattcaaatatgtatc	3676
Db	2566	ggggaaaatgtgcgcggaaacccctatttgtttatttttctaaatacattcaaatatgtatc	2625
QY	3677	cgctcatgagacaataaacctgataaatgcttcaataataattgaaaaaggaaagagtatga	3736
Db	2626	cgctcatgagacaataaacctgataaatgcttcaataataattgaaaaaggaaagagtatga	2685
QY	3737	gtattcaacatttcgctgtgcgcccttatcccttttttgcggcatttttgcccttcctgttt	3796
Db	2686	gtattcaacatttcgctgtgcgcccttatcccttttttgcggcatttttgcccttcctgttt	2745
QY	3797	ttgctcacccagaaaacgctggtgaaaagttaaagatgctgaagatcagttgggtgcacgag	3856
Db	2746	ttgctcacccagaaaacgctggtgaaaagttaaagatgctgaagatcagttgggtgcacgag	2805
QY	3857	tgggttacatcgaactggatctcaacagcggtaagatccttgagagtttttcgccccgaag	3916
Db	2806	tgggttacatcgaactggatctcaacagcggtaagatccttgagagtttttcgccccgaag	2865

QY 3917 aacgttttccaatgatgagcactttttaagtctctgctatgtggcgcggtattatcccgtta 3976
Db 2866 aacgttttccaatgatgagcactttttaagtctctgctatgtggcgcggtattatcccgtta 2925
QY 3977 ttgacgcgcgggcaagcaactcggtcgcgcgacatacactattctcagaatgacttggttg 4036
Db 2926 ttgacgcgcgggcaagcaactcggtcgcgcgacatacactattctcagaatgacttggttg 2985
QY 4037 agtactcaccagtcacagaaaaagcatcttacggatggcgcaacttactctgacaacgatcggag 4096
Db 2986 agtactcaccagtcacagaaaaagcatcttacggatggcgcaacttactctgacaacgatcggag 3045
QY 4097 gtgctgccataaaccatgagtataaacactgcggccaaacttactctgacaacgatcggag 4156
Db 3046 gtgctgccataaaccatgagtataaacactgcggccaaacttactctgacaacgatcggag 3105
QY 4157 gaccgaaggagtaacgccttttttgacaaacatgggggatcatgtaaactcgccttgatc 4216
Db 3106 gaccgaaggagtaacgccttttttgacaaacatgggggatcatgtaaactcgccttgatc 3165
QY 4217 gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctg 4276
Db 3166 gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctg 3225
QY 4277 tagcaatggcaaacagttgcgcaaaactattaaactggcgaaactacttactctagcttccc 4336
Db 3226 tagcaatggcaaacagttgcgcaaaactattaaactggcgaaactacttactctagcttccc 3285
QY 4337 ggcaacaattaatagactggatggagcgggataaagtgcaggaccacttctgcgctcgg 4396
Db 3286 ggcaacaattaatagactggatggagcgggataaagtgcaggaccacttctgcgctcgg 3345
QY 4397 cccctccggctggctggtttattgctgataaaatctggagccggtgagcgtgggtctcgcg 4456
Db 3346 cccctccggctggctggtttattgctgataaaatctggagccggtgagcgtgggtctcgcg 3405
QY 4457 gtatcattgcagcactgggcccagatggttaagccctcccgatcgtagttatctacacga 4516
Db 3406 gtatcattgcagcactgggcccagatggttaagccctcccgatcgtagttatctacacga 3465
QY 4517 cggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagtgccctcac 4576
Db 3466 cggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagtgccctcac 3525
QY 4577 tgattaaagcattggaactgtcagaccaaagtttactcaatatatacttttagattgatttaa 4636
Db 3526 tgattaaagcattggaactgtcagaccaaagtttactcaatatatacttttagattgatttaa 3585
QY 4637 aacttcatttttaatttaaaaggatctaggtgaagatcccttttttgataatctcatgacca 4696
Db 3586 aacttcatttttaatttaaaaggatctaggtgaagatcccttttttgataatctcatgacca 3645
QY 4697 aaatcccttaacgtgaggttttctgtccactgagcgtcagaccccgtagaaaaagatcaaaag 4756
Db 3646 aaatcccttaacgtgaggttttctgtccactgagcgtcagaccccgtagaaaaagatcaaaag 3705
QY 4757 gatctcttgagatcccttttttctgcgcgtaatctgctgcttgcaaaacaaaaaacccac 4816
Db 3706 gatctcttgagatcccttttttctgcgcgtaatctgctgcttgcaaaacaaaaaacccac 3765
QY 4817 cgctaccagcgggtggttggtttgcgggatcaagagctaaccaactctttttccgaaggtaa 4876
Db 3766 cgctaccagcgggtggttggtttgcgggatcaagagctaaccaactctttttccgaaggtaa 3825
QY 4877 ctggcttcagcagagcgcagataccaaatactgtccttctagtgtagcgtagtgagcc 4936
Db 3826 ctggcttcagcagagcgcagataccaaatactgtccttctagtgtagcgtagtgagcc 3885
QY 4937 accacttcaagaactctgtagcaccgcctacatacctcgtctgctgctaactcctgttaccag 4996
Db 3886 accacttcaagaactctgtagcaccgcctacatacctcgtctgctgctaactcctgttaccag 3945
QY 4997 tggctgctgccagtggcgataaagtcgtgtcttaccgggttggaactcaagacgatagttac 5056

Db 3946 tggctgctgccagtggcgataaagtcgtgtcttaccgggttggaactcaagacgatagttac 4005
QY 5057 cggataaggcgagcgggtcgggctgaacgggggttcgtgcacacagcccagcttggagc 5116
Db 4006 cggataaggcgagcgggtcgggctgaacgggggttcgtgcacacagcccagcttggagc 4065
QY 5117 gaacgacctacacccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttc 5176
Db 4066 gaacgacctacacccgaactgagatacctacagcgtgagcattgagaaagcgccacgcttc 4125
QY 5177 ccgaagggagaaaaagcggaacaggtatcccggtaaagcggcaggggtcggaacaggagagcgca 5236
Db 4126 ccgaagggagaaaaagcggaacaggtatcccggtaaagcggcaggggtcggaacaggagagcgca 4185
QY 5237 cgagggagcttccaggggaaaaacgcctggtatctttatagtcctgtcgggttccgcacc 5296
Db 4186 cgagggagcttccaggggaaaaacgcctggtatctttatagtcctgtcgggttccgcacc 4245
QY 5297 tctgacttgagcgtcgatttttctgctgctcgtcagggggcggaacctatggaaaaacg 5356
Db 4246 tctgacttgagcgtcgatttttctgctgctcgtcagggggcggaacctatggaaaaacg 4305
QY 5357 ccagcaacgcggccctttttacgggttcctggcccttttgcctggcccttttgcctcacatgttct 5416
Db 4306 ccagcaacgcggccctttttacgggttcctggcccttttgcctggcccttttgcctcacatgttct 4365
QY 5417 ttcctgcgttatccccctgattctgtggataaacctgattaccgccttttgagtgagctgata 5476
Db 4366 ttcctgcgttatccccctgattctgtggataaacctgattaccgccttttgagtgagctgata 4425
QY 5477 ccgctcgcgcagccggaacgaccgagcgcagcagtcagtgagcaggaagcggaagagc 5536
Db 4426 ccgctcgcgcagccggaacgaccgagcgcagcagtcagtgagcaggaagcggaagagc 4485
QY 5537 gcccaatacgcgaacccctctcccgcgcgttggcgattcattaaatgtagttagctca 5596
Db 4486 gcccaatacgcgaacccctctcccgcgcgttggcgattcattaaatgtagttagctca 4545
QY 5597 acaggtttcccgactggaaagcgggagtcagtgagcgcgaacgaatgaatgtgagttagctca 5656
Db 4546 acaggtttcccgactggaaagcgggagtcagtgagcgcgaacgaatgaatgtgagttagctca 4605
QY 5657 ctcataggcacccccaggttttacactttatgcttccggcctgctgattgtgtggaattg 5716
Db 4606 ctcataggcacccccaggttttacactttatgcttccggctcgtatgtgtgtggaattg 4665
QY 5717 tgagcggataacaatttcacacaggaacacagctatgacctgattacg 5764
Db 4666 tgagcggataacaatttcacacaggaacacagctatgacctgattacg 4713

RESULT 13
US-09-194-285-8
; Sequence 8, Application US/09194285
; Patent No. 6355479
; GENERAL INFORMATION:
; APPLICANT: Webb, Susan R.
; APPLICANT: Winqvist, Ola
; APPLICANT: Karlsson, Lars
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; TITLE OF INVENTION: MHC Class II Antigen Presenting Systems
; TITLE OF INVENTION: and Methods for Activating CD4+ T Cells
; FILE REFERENCE: TSRI 536.1
; CURRENT APPLICATION NUMBER: US/09/194,285
; CURRENT FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: PCT/US97/08697
; PRIOR FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: US 60/018,175
; PRIOR FILING DATE: 1996-05-23
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 8
; LENGTH: 4724
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-194-285-8

Query Match 45.5%; Score 2624.8; DB 4; Length 4724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2626; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	3137	cactggcgctgcttttacaaacgctgctgactgggaaacccctggcggttacccaacttaatc	3196
Db	2097	cactggcgctgcttttacaaacgctgctgactgggaaacccctggcggttacccaacttaatc	2156
QY	3197	gccttgcagcacatcccccttttcgcccagctggcgtaatagcgaagaggccgcaccgac	3256
Db	2157	gccttgcagcacatcccccttttcgcccagctggcgtaatagcgaagaggccgcaccgac	2216
QY	3257	gcccttcccaacagttgcgcagccctgaatggcgaaatggcgccctgatacgcggtattttctcc	3316
Db	2217	gcccttcccaacagttgcgcagccctgaatggcgaaatggcgccctgatacgcggtattttctcc	2276
QY	3317	ttacgcattctgctcggtattttcacaccgcgcatatggtgcactctctcagtaacaatctgctctg	3376
Db	2277	ttacgcattctgctcggtattttcacaccgcgcatatggtgcactctctcagtaacaatctgctctg	2336
QY	3377	atgccgcatagttaagccagccccgcacaccccgccaacaccccgctgacgcgccctgacggg	3436
Db	2337	atgccgcatagttaagccagccccgcacaccccgccaacaccccgctgacgcgccctgacggg	2396
QY	3437	cttgctctgctcccgccatcccgctttacagacaagctgtgaccgtctccgggagctgcatgt	3496
Db	2397	cttgctctgctcccgccatcccgctttacagacaagctgtgaccgtctccgggagctgcatgt	2456
QY	3497	gtcagaggttttcaccgctcatcacccgaacgcgcgcagacgaagggcctcgtgatacgcc	3556
Db	2457	gtcagaggttttcaccgctcatcacccgaacgcgcgcagacgaagggcctcgtgatacgcc	2516
QY	3557	tatttttataaggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttc	3616
Db	2517	tatttttataaggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttc	2576
QY	3617	ggggaaatgtgcgcggaaacccctatttgtttattttctaaataacattcaaatatgtatc	3676
Db	2577	ggggaaatgtgcgcggaaacccctatttgtttattttctaaataacattcaaatatgtatc	2636
QY	3677	cgctcatagagacaataaacctgatataatgcttcaataattgaaaaaggaaagatata	3736
Db	2637	cgctcatagagacaataaacctgatataatgcttcaataattgaaaaaggaaagatata	2696
QY	3737	gtattcaacatttccgtgtcgcccttattcccttttttgcggcatttttgccttccctgttt	3796
Db	2697	gtattcaacatttccgtgtcgcccttattcccttttttgcggcatttttgccttccctgttt	2756
QY	3797	ttgctcaccagaaacgcgtggtgaaagtataaagatgctgaagatcagttgggtgcacgag	3856
Db	2757	ttgctcaccagaaacgcgtggtgaaagtataaagatgctgaagatcagttgggtgcacgag	2816
QY	3857	tggtttacatcgaaactggatctcaacagcggtaagatcccttgagagtttttcgccccgaag	3916
Db	2817	tggtttacatcgaaactggatctcaacagcggtaagatcccttgagagtttttcgccccgaag	2876
QY	3917	aacgttttccaatgatgagcacatttttaagtctgctatgtggcgcggtattatcccgtta	3976
Db	2877	aacgttttccaatgatgagcacatttttaagtctgctatgtggcgcggtattatcccgtta	2936
QY	3977	ttgacgcggggcaagagcaactcggctccgcgcatacacactattctcagaatgacttggttg	4036
Db	2937	ttgacgcggggcaagagcaactcggctccgcgcatacacactattctcagaatgacttggttg	2996
QY	4037	agtactcacagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgca	4096

Db	2997	agtactcaccagtcacagaaaaagcatctttacggatggcatgacagtaagagaattatgca	3056
QY	4097	gtgctgccataaaccatgagtataaacactgcggccaaacttacttcttgacaacgatacggag	4156
Db	3057	gtgctgccataaaccatgagtataaacactgcggccaaacttacttcttgacaacgatacggag	3116
QY	4157	gaccgaaggaggtaaaccgctttttttgcacaacatgggggatcatgttaactcgccttgatc	4216
Db	3117	gaccgaaggaggtaaaccgctttttttgcacaacatgggggatcatgttaactcgccttgatc	3176
QY	4217	gttgggaaccggagctgaatgaagccataccaaaacgacgagcgtgacacacgatagcctg	4276
Db	3177	gttgggaaccggagctgaatgaagccataccaaaacgacgagcgtgacacacgatagcctg	3236
QY	4277	tagcaatggcaacaaactgtgcgcaaaactattaaactgggaactacttacttagcttccc	4336
Db	3237	tagcaatggcaacaaactgtgcgcaaaactattaaactgggaactacttacttagcttccc	3296
QY	4337	ggcaacaatttaatagactggatggaggcgataaaagtgcaggaccacttctcgctcgg	4396
Db	3297	ggcaacaatttaatagactggatggaggcgataaaagtgcaggaccacttctcgctcgg	3356
QY	4397	cccttcccggtcggtgtttattgctgataaaatctggagccggtgagcgtgggtctctcgcg	4456
Db	3357	cccttcccggtcggtgtttattgctgataaaatctggagccggtgagcgtgggtctctcgcg	3416
QY	4457	gtatcattgcagcactggggccagatggtaagccctcccgtatcgttagttatctacacga	4516
Db	3417	gtatcattgcagcactggggccagatggtaagccctcccgtatcgttagttatctacacga	3476
QY	4517	cggggagtcaaggcaactatggatgaacgaatagacagatcgcgtgagataggtgcctcac	4576
Db	3477	cggggagtcaaggcaactatggatgaacgaatagacagatcgcgtgagataggtgcctcac	3536
QY	4577	tgattaaagcatttggttaactgtcagaccaaagtttactcatatatacttttagattttaa	4636
Db	3537	tgattaaagcatttggttaactgtcagaccaaagtttactcatatatacttttagattttaa	3596
QY	4637	aacttcattttttaatttaaaggatctagtggaagatcccttttttgataatctcatgacca	4696
Db	3597	aacttcattttttaatttaaaggatctagtggaagatcccttttttgataatctcatgacca	3656
QY	4697	aaatcccttaacgtagtatttctgttccactgcggtcagaccccgtagaaaaagatcaaa	4756
Db	3657	aaatcccttaacgtagtatttctgttccactgcggtcagaccccgtagaaaaagatcaaa	3716
QY	4757	gatcttcttgagatccctttttctgcgcgttaactgtctgtcgttgcaaaacaaaaaacac	4816
Db	3717	gatcttcttgagatccctttttctgcgcgttaactgtctgtcgttgcaaaacaaaaaacac	3776
QY	4817	cgctaccagcgggtgtgttgcgggatacaagagctaccacactcttttccggaaggttaa	4876
Db	3777	cgctaccagcgggtgtgttgcgggatacaagagctaccacactcttttccggaaggttaa	3836
QY	4877	ctggcttcagcagagcgcagatataccaaatactgttccttcttagtgcgttagttaggcc	4936
Db	3837	ctggcttcagcagagcgcagatataccaaatactgttccttcttagtgcgttagttaggcc	3896
QY	4937	accacttcaaagaactctgttagcaccgcctacatacctcgtctgtctaactcctgttaccag	4996
Db	3897	accacttcaaagaactctgttagcaccgcctacatacctcgtctgtctaactcctgttaccag	3956
QY	4997	tggtcgtgccagtgccgataagtcgttcttaccgggttggaactcaagacgatagttac	5056
Db	3957	tggtcgtgccagtgccgataagtcgttcttaccgggttggaactcaagacgatagttac	4016
QY	5057	cggataaggcgagcgggtcggtgaacgggggggttcgtgcacacagcccagcttggagc	5116
Db	4017	cggataaggcgagcgggtcggtgaacgggggggttcgtgcacacagcccagcttggagc	4076
QY	5117	gaacgacctacacccgaactgagatacctacagcgtgagctatgagaaaagccacgcttc	5176
Db	4077	gaacgacctacacccgaactgagatacctacagcgtgagctatgagaaaagccacgcttc	4136

QY 5177 ccgaaggagaaagcgagcagggtatccggtaagcggcagggctcggaacagagagcgca 5236
Db 4137 ccgaaggagaaagcgagcagggtatccggtaagcggcagggctcggaacagagagcgca 4196
QY 5237 cgaggagcttccaggggaaacgcctgggtatctttatagtcctgtcgggttcgccacc 5296
Db 4197 cgaggagcttccaggggaaacgcctgggtatctttatagtcctgtcgggttcgccacc 4256
QY 5297 tctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatggaaaaacg 5356
Db 4257 tctgacttgagcgtcgatttttgtgatgctcgtcagggggcgagcctatggaaaaacg 4316
QY 5357 ccagcaacgcggcccttttacgggttcctggccttttgcggccttttgcacatgttct 5416
Db 4317 ccagcaacgcggcccttttacgggttcctggccttttgcggccttttgcacatgttct 4376
QY 5417 ttccctgcgttatccctgattcttgtggataaccgtattaccgcctttgagtgagctgata 5476
Db 4377 ttccctgcgttatccctgattcttgtggataaccgtattaccgcctttgagtgagctgata 4436
QY 5477 ccgctcgcgcagcgaacgacgagcgcagcagcagtcagtgagcaggaagcagc 5536
Db 4437 ccgctcgcgcagcgaacgacgagcgcagcagtcagtgagcaggaagcagc 4496
QY 5537 gcccaatcgcgaacgcctctcccgcgcttgcccgattcattaatgagctggcacg 5596
Db 4497 gcccaatcgcgaacgcctctcccgcgcttgcccgattcattaatgagctggcacg 4556
QY 5597 acaggtttcccgactggaaagcggcagtgagcgaacgcaattaatgagttagctca 5656
Db 4557 acaggtttcccgactggaaagcggcagtgagcgaacgcaattaatgagttagctca 4616
QY 5657 ctcattaggcaccacaggcctttacacctttatgcttccggctcgtatgttgtggaattg 5716
Db 4617 ctcattaggcaccacaggcctttacacctttatgcttccggctcgtatgttgtggaattg 4676
QY 5717 tgagcggataacaatttcacacaggaacacagctatgaccatgattacg 5764
Db 4677 tgagcggataacaatttcacacaggaacacagctatgaccatgattacg 4724

RESULT 14
US-08-844-274-20
; Sequence 20, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-08-844-274-20

Query Match 45.5%; Score 2623.2; DB 4; Length 7560;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2625; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3137 cactggccgctcgttttacaaacgctcgtgactgggaaaaacccctggcgttacccaacttaatc 3196
Db 4766 cactggccgctcgttttacaaacgctcgtgactgggaaaaacccctggcgttacccaacttaatc 4825
QY 3197 gccttgcagcacatcccccttttcgccagcctggcgtaataagcgaagagggccgcgacccgac 3256
Db 4826 gccttgcagcacatcccccttttcgccagcctggcgtaataagcgaagagggccgcgacccgac 4885
QY 3257 gcccttcccaaacagttgcgcagccctgaaatggcgaaatggcgccctgatgcggtattttctcc 3316
Db 4886 gcccttcccaaacagttgcgcagccctgaaatggcgaaatggcgccctgatgcggtattttctcc 4945
QY 3317 ttacgcatctgtcggtattttcacacccgcataatggcgactctcagtcacatactgtctctg 3376
Db 4946 ttacgcatctgtcggtattttcacacccgcataatggcgactctcagtcacatactgtctctg 5005
QY 3377 atgccgcatagtttaagccagccccgcacacccgcgaacacccgcctgacgcgcctgacggg 3436
Db 5006 atgccgcatagtttaagccagccccgcacacccgcgaacacccgcctgacgcgcctgacggg 5065
QY 3437 cttgtctgctcccgcatccgccttacagacaagctgtgacccgtctccgggagctgcatgt 3496
Db 5066 cttgtctgctcccgcatccgccttacagacaagctgtgacccgtctccgggagctgcatgt 5125
QY 3497 gtcagaggtttttcacccgtcatcacggaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3556
Db 5126 gtcagaggtttttcacccgtcatcacggaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5185
QY 3557 tattttataggttaatgcatgataataatgtttatttttctaaatacatattcaaatatgtatc 3616
Db 5186 tattttataggttaatgcatgataataatgtttatttttctaaatacatattcaaatatgtatc 5245
QY 3617 ggggaaatgtgcgcggaacccctatttgtttatttttctaaatacatattcaaatatgtatc 3676
Db 5246 ggggaaatgtgcgcggaacccctatttgtttatttttctaaatacatattcaaatatgtatc 5305
QY 3677 cgctcatgagacataaaccttgataaaatgcttcaataatattgaaaaaggaagagtatga 3736
Db 5306 cgctcatgagacataaaccttgataaaatgcttcaataatattgaaaaaggaagagtatga 5365
QY 3737 gtattcaacatttcctgctgcgccttattcccttttttgcggcatttttgcccttccctgttt 3796
Db 5366 gtattcaacatttcctgctgcgccttattcccttttttgcggcatttttgcccttccctgttt 5425
QY 3797 ttgctcaccacagaaacgctggtgaaaagtataaagatgctgaagatcagttgggtgcacgag 3856
Db 5426 ttgctcaccacagaaacgctggtgaaaagtataaagatgctgaagatcagttgggtgcacgag 5485
QY 3857 tgggttacatcgaactggatctcaacagcggtaagatcccttgagagttttcgccccgaag 3916
Db 5486 tgggttacatcgaactggatctcaacagcggtaagatcccttgagagttttcgccccgaag 5545
QY 3917 aacgttttccaaatgatgagcacattttaagttctgctatgtggcgcggtattatcccgtta 3976
Db 5546 aacgttttccaaatgatgagcacattttaagttctgctatgtggcgcggtattatcccgtta 5605
QY 3977 ttgacgcggggcaagagcaactcggtcgcgcgcatacacactatttctcagaatgacttggttg 4036
Db 5606 ttgacgcggggcaagagcaactcggtcgcgcgcatacacactatttctcagaatgacttggttg 5665
QY 4037 agtactcaccagtcacagaaaaagcatctttacggatggcatgacagtaagagaattatgca 4096
Db 5666 agtactcaccagtcacagaaaaagcatctttacggatggcatgacagtaagagaattatgca 5725
QY 4097 gtgctgccataaaccatgagtataaacactcgggccaaacttacttctgacaacgacgcggag 4156
Db 5726 gtgctgccataaaccatgagtataaacactcgggccaaacttacttctgacaacgacgcggag 5785
QY 4157 gaccgaaggagctaaccgccttttttgacaaacatgggggatcatgttaactgccttgatc 4216
Db 5786 gaccgaaggagctaaccgccttttttgacaaacatgggggatcatgttaactgccttgatc 5845
QY 4217 gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacacacgacgatgcctg 4276

Db 5846 gttaggaaccgctgaatgaagccataccaaacgacgagcgtgacaccagatgcctg 5905
QY 4277 tagcaatggcaacaacgttgccaaactattaaactggcgaaactactactctagctccc 4336
Db 5906 tagcaatggcaacaacgttgccaaactattaaactggcgaaactactactctagctccc 5965
QY 4337 ggcaacaattaatagactggatggaggcggaataaaagtgcaggaccacttctgcgctcgg 4396
Db 5966 ggcaacaattaatagactggatggaggcggaataaaagtgcaggaccacttctgcgctcgg 6025
QY 4397 ccttccggctggctggtttattgtctgataaaatctggagccggtgagcgtgggtctcgcg 4456
Db 6026 ccttccggctggctggtttattgtctgataaaatctggagccggtgagcgtgggtctcgcg 6085
QY 4457 gtatcattgcagcactggggccagatggtaagccctcccgtagttagttatctacacga 4516
Db 6086 gtatcattgcagcactggggccagatggtaagccctcccgtagttagttatctacacga 6145
QY 4517 cggggagtcaggcaactatggatgaacgaataagacagatcgcgtgagataggtgcctcac 4576
Db 6146 cggggagtcaggcaactatggatgaacgaataagacagatcgcgtgagataggtgcctcac 6205
QY 4577 tgattaagcattggtaaactgtcagaccaagttagtctatatactatcttagattgattaa 4636
Db 6206 tgattaagcattggtaaactgtcagaccaagttagtctatatactatcttagattgattaa 6265
QY 4637 aacttcatttttaatttaaaggatctaggtgaagatcccttttgataatctcatgacca 4696
Db 6266 aacttcatttttaatttaaaggatctaggtgaagatcccttttgataatctcatgacca 6325
QY 4697 aaatcccttaacgtgagttttgttccactgagcgtcagaccccgtagaaaaagatcaaaag 4756
Db 6326 aaatcccttaacgtgagttttgttccactgagcgtcagaccccgtagaaaaagatcaaaag 6385
QY 4757 gatcttcttgagatcccttttttctgcgcgtaactctgctgttgcaaaacaaaaaacac 4816
Db 6386 gatcttcttgagatcccttttttctgcgcgtaactctgctgttgcaaaacaaaaaacac 6445
QY 4817 cgctaccagcggtggtttgttgccggatcaagagctacaaactcttttccgaaggttaa 4876
Db 6446 cgctaccagcggtggtttgttgccggatcaagagctacaaactcttttccgaaggttaa 6505
QY 4877 ctgcttcagcagcagcagcagataccaaatactgtccttctagtgtagccgtagtgcc 4936
Db 6506 ctgcttcagcagcagcagcagataccaaatactgtccttctagtgtagccgtagtgcc 6565
QY 4937 accacttcaagaactctgtagcaccgcctacatacctcgtctgtgtaactcctgttaccag 4996
Db 6566 accacttcaagaactctgtagcaccgcctacatacctcgtcgtgtaactcctgttaccag 6625
QY 4997 tggctgctgccagtgccgataaagtctgtcttaccgggttggaactcaagacgatatgttac 5056
Db 6626 tggctgctgccagtgccgataaagtctgtcttaccgggttggaactcaagacgatatgttac 6685
QY 5057 cggataaaggcgcagcgggtcggtgaacgggggttcgtgcacacagcccagcttgagc 5116
Db 6686 cggataaaggcgcagcgggtcggtgaacgggggttcgtgcacacagcccagcttgagc 6745
QY 5117 gaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagccacgcttc 5176
Db 6746 gaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagccacgcttc 6805
QY 5177 ccgaaggagaaaggcggacaggtatccggtaagcggcgaggttcggaacaggagagcgca 5236
Db 6806 ccgaaggagaaaggcggacaggtatccggtaagcggcgaggttcggaacaggagagcgca 6865
QY 5237 cgaaggagcttccagggggaaaacgcctggtatctttatagctcctgtcgggtttcgccacc 5296
Db 6866 cgaaggagcttccagggggaaaacgcctggtatctttatagctcctgtcgggtttcgccacc 6925
QY 5297 tctgacttgagcgtcgatcttctgtgatgctcgtcagggggcgagcctatggaaaaacg 5356
|||||

Db 6926 tctgacttgagcgtcgatctcgattttttgtgatgctcgtcagggggcgagcctatggaaaaacg 6985
QY 5357 ccagcaacgcggcccttttttaacggttccctggcccttttgcggccttttgcctcacatgttct 5416
Db 6986 ccagcaacgcggcccttttttaacggttccctggcccttttgcggccttttgcctcacatgttct 7045
QY 5417 ttctgcggttatccctgattctgtggataaacggtattaccgcctttgagtgagctgata 5476
Db 7046 ttctgcggttatccctgattctgtggataaacggtattaccgcctttgagtgagctgata 7105
QY 5477 ccgctgcgcgagccgaacgacgagcgcagcagcagtgagtgagtgaggaagcgggaagc 5536
Db 7106 ccgctgcgcgagccgaacgacgagcgcagcagtgagtgagtgaggaagcgggaagc 7165
QY 5537 gcccaatacgcaaacccgctctcccgcgcttgccgattcatttaattgagctggcacg 5596
Db 7166 gcccaatacgcaaacccgctctcccgcgcttgccgattcatttaattgagctggcacg 7225
QY 5597 acaggtttcccgactggaaagcgggcagtgagcgaacgcaacgcaattagtgagctca 5656
Db 7226 acaggtttcccgactggaaagcgggcagtgagcgaacgcaattagtgagctca 7285
QY 5657 ctcattaggcacccccaggctttacactttatgcttccggctcgtatgtgtggaattg 5716
Db 7286 ctcattaggcacccccaggctttacactttatgcttccggctcgtatgtgtggaattg 7345
QY 5717 tgagcggaataacaatttcacacagagaaacagctatgacctgattacg 5764
Db 7346 tgagcggaataacaatttcacacagagaaacagctatgacctgattacg 7393

RESULT 15
US-08-385-335A-8
; Sequence 8, Application US/08385335A
; Patent No. 5891680
; GENERAL INFORMATION:
; APPLICANT: Lieschke, Graham J.
; APPLICANT: Mulligan, Richard C.
; TITLE OF INVENTION: Bioactive Fusion Proteins
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,335A
; FILING DATE: 08-FEB-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI95-01
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-385-335A-8

Query Match		45.4%;	Score 2620.8;	DB 2;	Length 6350;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 2633;		Conservative	0;	Mismatches	2;
				Indels	1;
				Gaps	1;
Qy	3132	gaattcactggcgcgtcgctttttacaacgctgctgactgggaaaaaccctggcgttaccccaact	3191		
Db	3715	GAATTCACCTGGCCGTCGTTTACAAACGCTGCTGACTGGGAAAAACCCCTGGCGTTACCCCAACT	3774		
Qy	3192	taatcgccctgcagcacacatccccctttcgccagctggcgttaataagagaagagcccgac	3251		
Db	3775	TAATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCAAGAGGCCCGCAC	3834		
Qy	3252	cgatcgcccttcccacaagttgcgcagctgaatggcgaaatggcgctgatcggtattt	3311		
Db	3835	CGATCGCCCTTCCCAACAGTTTGGCAGCGCTGAATGGCGAATGGCGCTGATCGGTAATT	3894		
Qy	3312	tctccttacgcatctgtgcggtattttcacaccgcataatggtgcactctcagtacaatctg	3371		
Db	3895	TCCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACCTCAGTACAATCTG	3954		
Qy	3372	ctctgatcgcgcatagttaagccagccccgcagaccccccaacacccgcgtgacgcgccctg	3431		
Db	3955	CTCTGATGCCCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCCCCCTG	4014		
Qy	3432	acgggcttgcgtcctccggcgtacgttacagacaagctgtgacgcgtctccgggagctg	3491		
Db	4015	ACGGGCTTGTCTCCCGGCATCCCGTTACAGACAAGCTGTGACCGTCTCCGGGAGCTG	4074		
Qy	3492	catgtgtcagagggttttcaccgctcatcacggaaaacgcgcga-gacgaaaaggccctcgtga	3550		
Db	4075	CATGTGTACAGAGGTTTTCACCGTCTACCCGATCACCAGAAACGCCGATGACGAAAGGGCCTCGTGA	4134		
Qy	3551	tacgcctattttataggttaatgtcatgataataatggtttcttagacgctcaggtggca	3610		
Db	4135	TACGCCTATTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTAGACGTCAGGTGGCA	4194		
Qy	3611	cttttcggggaatgtgcgcggaacccctatttgtttatttttctaaatacattcaata	3670		
Db	4195	CTTTTCGGGGAATGTGCGGGAACCCCTATTGTGTTATTTTCTAAATACATTCAAATA	4254		
Qy	3671	tgtatccgctcatgagacaataaaccttgataaatgcttcaataataattgaaaaaggaaga	3730		
Db	4255	TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA	4314		
Qy	3731	gtatgagtattcaacatttcgcgtgtcgccttattcccttttttgcggcattttgccttc	3790		
Db	4315	GTATGAGTATTCAACATTTCCGTTGTCGCCCTTATTCCTTTTTTTTGGCGCATTTTGGCCTTC	4374		
Qy	3791	ctgttttttgcacccagaaaacgctggtgaaaagtaaaagatgctgaagatcagttgggtg	3850		
Db	4375	CTGTTTGTCTACCCAGAAACGCTGGTGAAGATAAAGATGCTGAAGATCAGTTGGGTG	4434		
Qy	3851	cacgagtgggttacatcgaactggatcgaacagcggtgaagatccttgagagttttcgcc	3910		
Db	4435	CACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC	4494		
Qy	3911	ccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcggtattat	3970		
Db	4495	CCGAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTATTAT	4554		
Qy	3971	cccgtattgacgcgggcaagagcaactcggctgcgcgcatacacactattctcagaatgact	4030		
Db	4555	CCCGTATTGACCGCGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAATGACT	4614		
Qy	4031	tggttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat	4090		
Db	4615	TGGTTGAGTACTACCCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAAT	4674		
Qy	4091	tatgcagtgctgcataaaccatgagtataacactgcggccaacttacttctgacaacga	4150		
Db	4675	TATGCAGTGTGCCATAAACCATGAGTGATAACACTCGGGCCAACCTTACTTCTGACRAACGA	4734		
Qy	4151	tcggaggaccegaaggagctaaccgctttttttgcacaacatgggggatcatgtaactcgcc	4210		

Db	4735	TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAACATGGGGGATCATGTAACCTCGCC	4794		
Qy	4211	ttgatcgttgggaaaccggagctgaaagaagccataccaaacgacgagcgtgacaccacga	4270		
Db	4795	TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCNAAACGACGAGCGTGACACCACGA	4854		
Qy	4271	tgccgttagcaatggcaacaacgttgcgcaaaactattaaactggcgaaactacttactctag	4330		
Db	4855	TGCCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGNACTACTTACTCTAG	4914		
Qy	4331	cttcccgcgcaacaattaatagactggatggaggcggaataaaagttagaggaccacttctgc	4390		
Db	4915	CTTCCCGGCAACAATTAAATAGACTGGATGGAGGGGATAAAAGTTGCAGGACCACCTTCTGC	4974		
Qy	4391	gctcggccctccggcgtggctggtttatttgctgataaaatctggagccggtgagcgtgggt	4450		
Db	4975	GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGGAGCCGCTGAGCGTGGGT	5034		
Qy	4451	ctcgcggtatcaattgcagcactggggccagatggtaagccctcccgctatcgtagttatct	4510		
Db	5035	CTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT	5094		
Qy	4511	acacgacggggagtcaggcaactatggatgaacgaataagacagatcgctgagataggtg	4570		
Db	5095	ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG	5154		
Qy	4571	cctcactgattaaagcattggtaactgtcagaccaagtttactcatatatactttagattg	4630		
Db	5155	CCTCACTGATTAAGCATTTGGTAACCTGTACAGACCAAGTTTACTCATATATACTTTAGATTG	5214		
Qy	4631	atttaaaactctcatcttttaatttaaaggatctaggtgaagatccctttttgataaatctca	4690		
Db	5215	ATTTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA	5274		
Qy	4691	tgacaaaaatcccttaacgtgagtttttcgttccactgagcgtcagaccccgtagaaaaa	4750		
Db	5275	TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGCTCAGACCCCGTAGAAAAGA	5334		
Qy	4751	tcaaaggatctcttgagatcccttttttctgcgcgtaactcgtcgttgcaaaacaaaa	4810		
Db	5335	TCAAAGGATCTTCTTGAGATCCTTTTCTGCGGTAATCTGCTGCTGCAAAACAAAA	5394		
Qy	4811	aaccacccgtaccagcggtggtttgtttgcgggatcaagagctaccaactctttttccga	4870		
Db	5395	AAACCACCGCTACAGCGGTGGTTGTTGTCGGGATCAAGAGCTACCAACTCTTTTTCGGA	5454		
Qy	4871	aggtaactggcttcagcagcagcagataccaaatactgtccttctagtgtagccgtagt	4930		
Db	5455	AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCCCTTCTAGTGTAGCCGTAGT	5514		
Qy	4931	taggccaccacttcaagaactctgtagcaccgcctacatacctcgtctctgctaaatcctgt	4990		
Db	5515	TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCCTACATACCTCGCTCTGCTAATCCTGT	5574		
Qy	4991	taccagtggctgctgccagtggcgataagtcgtgtcttaccgggttcggactcaagacgat	5050		
Db	5575	TACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTGGACTCAAGACGAT	5634		
Qy	5051	agttaccggataaggcgagcgtcggggtgaaacgggggttcctgctcacacagcccagct	5110		
Db	5635	AGTTACCGGATAAAGGCGCAGCGGTCCGGGTGAACGGGGGTTTCGTGCACACAGCCCCAGCT	5694		
Qy	5111	tggagcgaacgacctacaccgaactgagatcacctacagcgtgagctatgagaaagcgcca	5170		
Db	5695	TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAAGCGCCA	5754		
Qy	5171	cgcttcccgaaggagaaaggcgacaggtatccggtaagcggcgagcgttcggaacaggag	5230		
Db	5755	CGCTTCCCGAAGGGAGAAAGGCGGACAGTATCCGGTAAGCGGCAGGGTCGGAACAGGAG	5814		
Qy	5231	agcgcacgaggggagcttccagggggaacgccttggtatctttatagtcctgctgggttttc	5290		

Db 5815 AGCGACAGGGAGGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCCTGTGCGGTTTC 5874

QY 5291 gccacctctgacttgagcgctcgatttttbtgtgatgctcgtcaggggggagccctatgga 5350
|||||

Db 5875 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA 5934

QY 5351 aaaagccagcaacgcgcccttttttacgggttcctggcccttttgctggccttttgctcaca 5410
|||||

Db 5935 AAAAGCCAGCAACGCGGCCCTTTTACGGTTCTCTGGCCCTTTTGCTGGCCTTTTGCTCACA 5994

QY 5411 tgttctttcctgcgttatccccctgattctgtggataaacgtattaccgcctttgagtgag 5470
|||||

Db 5995 TGTTCTTCTCCTGCGTTATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG 6054

QY 5471 ctgataccgctcgccgcagccgaacgcagccagcgagtcagtgagcgaggaagcgg 5530
|||||

Db 6055 CTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGGAAGCGG 6114

QY 5531 aagagcgcccaatacgcgaacgcctctctcccgcgcttggtggcgattcattaatgcagct 5590
|||||

Db 6115 AAGAGGCCCAATACGCAAACCGCCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCT 6174

QY 5591 ggcagcacaggtttcccgactgaaagcgggcagtgagcgaacgcaattaatgtgagtt 5650
|||||

Db 6175 GGCACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAAATGTGAGTT 6234

QY 5651 agctcactcattagggcacccccaggcctttacacctttatgcttccggctcgtatgtgtg 5710
|||||

Db 6235 AGCTCACTCATTAGGCACCCCGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 6294

QY 5711 gaattgtgagcgggataaacaatttcacacagggaacagctatgacctgattacgcc 5766
|||||

Db 6295 GAATTGTGCGGGATACAAATTTACACAGGAACAGCTATGACCATGATTACGCC 6350

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:10:48 ; Search time 9324.87 Seconds
(without alignments)
3871.183 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atgaggccccgcagtgctt.....aatgtcagcgccacctga 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:	
1:	gb_ba:*
2:	gb_htg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vi:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_vi:*
30:	em_htg_hum:*
31:	em_htg_inv:*
32:	em_htg_other:*
33:	em_htgo_inv:*

ALIGNMENTS

RESULT 1	
AX275256	
LOCUS	AX275256 1725 bp DNA linear PAT 29-OCT-2001
DEFINITION	Sequence 5 from Patent WO0171014.
ACCESSION	AX275256
VERSION	AX275256.1 GI:16547676
KEYWORDS	synthetic construct.
SOURCE	synthetic construct
ORGANISM	artificial sequence.
REFERENCE	1 (sites)
AUTHORS	Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE	Expression of recombinant human acetylcholinesterase in transgenic plants
JOURNAL	Patent: WO 0171014-A 5 27-SEP-2001; BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor, Tsafrir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason, Hugh S. (US)
FEATURES	Location/Qualifiers
source	1..1725 /organism="synthetic construct" /db_xref="taxon:32630" /note="synthetic human acetylcholinesterase gene optimized

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

		for expression in plants"			
BASE COUNT	273 a	553 c	571 g	328 t	
ORIGIN					
Query Match	100.0%;	Score 1725;	DB 6;	Length 1725;	
Best Local Similarity	100.0%;	Pred. No. 1e-294;			
Matches 1725;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps	
QY	1	atgagggccccgcagtgctgctgcacagccttcctcgcttccccactccttcctc	60		
Db	1	ATGAGGGCCCCCGCAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCACACTCCTTCTCTCTC	60		
QY	61	ctcctctggctcctgggtggaggagtgggggctgagggccggaggatgcagagctgctg	120		
Db	61	CTCCTCTGGCTCCTTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG	120		
QY	121	gtgacgggtgcgtggggccggctgcggggcattcgcctgaagacccccggggccctgctc	180		
Db	121	GTGACGGTGGCTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGGCCCTGTC	180		
QY	181	tctgctttcctgggcatccccctttgcggagggccaccccatgggaccccgctttctgcca	240		
Db	181	TCTGCTTTCCTGGGCATCCCCCTTTGCGAGCCACCCCATGGGACCCCGCTCGCTTCTTGCCA	240		
QY	241	ccggagcccaagcagccttggtcaggggtggttagacgctacaaccttcacagagtgtctgc	300		
Db	241	CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC	300		
QY	301	taccaatatgtggacacccctataccccagggttttgagggcaccgagatgtggaaccccaac	360		
Db	301	TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC	360		
QY	361	cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccggcctaca	420		
Db	361	CGTGAGCTGAGCGAGGAGACTGCCTGTACTCTAACGTGTGGACACCATACCCCCGGCCTACA	420		
QY	421	tcccccaacctgtcctcgtctggatctatgggggtggttcttctacagtggggccctcctcc	480		
Db	421	TCCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC	480		
QY	481	ttggacgtgtacgatggccgcttcttggtacagggccgagaggactgtgctggtgccatg	540		
Db	481	TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG	540		
QY	541	aactaccgggtgggagccctttggcttcttggccctgccggggagccgagagggccccgggc	600		
Db	541	AACTACCGGGTGGGAGCCTTTGGCTTCTCTGGCCCTGCCGGGGAGCCCGAGAGGCCCGGGC	600		
QY	601	aatgtgggtcctcctggatcagaggctggccctgcagtgggtgcaggagaacgtggcagcc	660		
Db	601	AATGTGGGTCTCCTTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC	660		
QY	661	ttcgggggtgaccccgacatcagtgacgtctgtttggggagagcgcgggagccgcctcggtg	720		
Db	661	TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCCGCCTCGGTG	720		
QY	721	ggcatgcacctgtgtccccgccccagccggggcctgttccacagggccgctgctgcagagc	780		
Db	721	GGCATGCACCTGTGTCCCCCGCCACAGCCGGGGCCTGTTCACAGGGCCGTGCTGCAGAGC	780		
QY	781	ggtgcccccaatggaccctggggccacgttgggcattgggagagggccccgtcgcagggccacg	840		
Db	781	GGTGCCCCCAATGGACCCCTGGGGCCACGCTGGGCATGGGAGAGGCCCTCGCAGGGGCCACG	840		
QY	841	cagctggccccaccttgtgggtgtcctccagggcggcactggtgggaatgacacagagctg	900		
Db	841	CAGCTGGCCCCACCTTGTGGGTGTCTCTCCAGGGGGCACTGGTGGGAATGACACAGAGCTG	900		
QY	901	gtagcctgccttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgtgctg	960		
Db	901	GTAGCCTGCCTTCGGACACGACGACGCGCAGGTCTCTGTTGAACCAACCAATGGCACGTGCTG	960		

Qy	961	cctcaagaaagcgtcttccggttctccttcctcgtgcctgtggtagatgagagacttcctcagt	1020
Db	961	CCTCAAGAAAGCGTCTTCCGGTTCTCTCTCGTTCCTGTGTTAGATGGAGACTTCTCTCAGT	1020
Qy	1021	gacacccccagaggccctcatcaacgcggggagagacttccacggcctcagaggtgctggtgggt	1080
Db	1021	GACACCCCAGAGGCCCTCATCAACGGGGGAGACTTCCACGGCCTGAGGTGCTGGTGGGT	1080
Qy	1081	gtggtgaagatgagggctcgtattttcttggttttacggggccccaggttcagcaaaagac	1140
Db	1081	GTGGTGAAAGATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCAGGCTTCAGCAAAGAC	1140
Qy	1141	aacgagtctctcatcagccggcgagttctcgtgctgacacagactggctgcatcccgag	1200
Db	1141	AACGAGTCTCTCATCAGCCGGCCGAGTTCCTGGCCGGGTGCGGGTCCGGGTTCCTCCAG	1200
Qy	1201	gtaagtacctggcagccgaggctgtggtcctgctattacacagactggctgcatcccgag	1260
Db	1201	GTAAGTGACCTGGCAGCCGAGGCTGTGTTCTGCTGCTTACACAGACTGGCTGCATCCCGAG	1260
Qy	1261	gacccggcacgcctgagggagggccctgagcgatgtggtgggcgacacaatgtcgtgtgc	1320
Db	1261	GACCCGGCACGCCTGAGGGAGGCCCTGAGCGATGTGGTGGGGACCAACAATGTCTGTGTC	1320
Qy	1321	cccgtggccagctggctggcgactggctgccccagggtgccccgggtcttacgcctacgtc	1380
Db	1321	CCCGTGGCCAGCTGGCTGGCGACTGGCTGCCCCAGGGTGCCCGGGTCTACGCCTACGTC	1380
Qy	1381	tttgaacacgctgcttccacgctctcctgccccctgtggatgggggtgccccacggctac	1440
Db	1381	TTTGAACACCGTGTCTCCACGCTCTCCTGGCCCCCTGTGGATGGGGTGCCCCACGGCTAC	1440
Qy	1441	gagatcgagttcatcttgggataccccctggaccctctctgaaactacacagggcagaggag	1500
Db	1441	GAGATCGAGTTCATCTTTGGGATCCCCCTGACCCCTCTCGAAACTACACGGCAGAGGAG	1500
Qy	1501	aaaatcttcgccccagcagctgatgcgatactggggccaaactttgcccgcacaggggacccc	1560
Db	1501	AAATCTTCGCCACGCGACTGATCGATACTTGGGCCAACTTTGCCCGCACAGGGGATCCC	1560
Qy	1561	aatgagccccgagagacccccaaaggccccacaaatggcccccgctacacggcgggctcagcag	1620
Db	1561	AATGAGCCCCGAGACCCCCAAGGCCCCACAAATGGCCCCCGTACACGGCGGGGCTCAGCAG	1620
Qy	1621	tacgttagctggacccctgcggccgctggaaggtgcggcgggggtgcggccccaggcctgc	1680
Db	1621	TACGTTAGTCTGGACCTGCGGCCGCTGGAGTGGCGGGGCTGCGGCCCCAGGCCTGC	1680
Qy	1681	gccttctggaaccgcttctcccccaaatgtctcagcgccacctga	1725
Db	1681	GCCTTCTGGAACCGCTTCTCTCCCCAAATTGCTCAGCGCCACCTGA	1725

RESULT 2

AR070206	LOCUS	AR070206	3096 bp	DNA	linear	PAT 18-FEB-2000
DEFINITION	Sequence 6 from patent US 5891725.					
ACCESSION	AR070206					
VERSION	AR070206.1	GI:7221094				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3096)					
AUTHORS	Soreg,H., Zakut,H. and Eckstein,F.					
TITLE	Synthetic antisense oligodeoxynucleotides and pharmaceutical compositions containing them					
JOURNAL	Patent: US 5891725-A 6 06-APR-1999;					
FEATURES	Location/Qualifiers					
source	1. .3096					
BASE COUNT	509 a 1089 c 872 g 626 t					
ORIGIN	/organism="unknown"					

Query Match 99.8%; Score 1722.4; DB 6; Length 3096;
Best Local Similarity 99.9%; Pred. No. 2.7e-294;
Matches 1723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	atgagccccccgcagtgctctgcacacgccttccctggcttccccactccttctcctc	60
Db	160	ATGAGCCCCCGCAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCCACTCTTCTCCTC	219
Qy	61	ctcctctggctcctcctgggtgagagagtgggggctgagggccgggaggtgcagagctgctg	120
Db	220	CTCCTCTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG	279
Qy	121	gtgacggtgcgtggggccggctgctggggcattcgccctgaagacccccccggggccctgtc	180
Db	280	GTGACGCTGCGTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCCGGGGCCCTGTC	339
Qy	181	tctgcttctcctgggcatacccccttggcgagccaccctatgggacccccctgccttctgcca	240
Db	340	TCTGCTTCTCTGGGCATCCCCCTTTGCGGAGCCACCCCATGGGACCCCTGCTTCTGCCA	399
Qy	241	ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttccagagtgtctgc	300
Db	400	CCGGAGCCCAAGCAGCCTTGCTCAGGGTGTAGACGCTACAACTTCCAGAGTGTCTGC	459
Qy	301	taccaatatgtgacaccctataccccagggttttgaggggcacagagatgtggaacccccaac	360
Db	460	TACCAATATGTGGACACCCTATACCAGGTTTGTAGGGCACCCAGATGTGGAACCCCAAC	519
Qy	361	cgtgagctgagcgaggactgcctgtacctcaacgtgtgggacacataccccccggcctaca	420
Db	520	CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCGGCCTACA	579
Qy	421	tccccacacccctgtcctcgtctcgtgatctatgggggtggcttctacagtgggcctcctcc	480
Db	580	TCCCCACCCCTGTCTCTGCTGTGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC	639
Qy	481	tggacgtgtacgatggccgcttcttggtacagggccgagagagactgtgctggtgccatg	540
Db	640	TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGTTCCATG	699
Qy	541	aactacggggtggagaccttggcttccctggccctgccggggagccgagagccccgggc	600
Db	700	AACtACCGGgTGGAGCCtTtTGGCTTCTTGCCCTGCCGGGAGCCGAGAGCCCCGGGC	759
Qy	601	aatgtgggtcctcctggatcagaggtggccctgcagtggggtgcaggagaaacgtggcagcc	660
Db	760	AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC	819
Qy	661	ttcgggggtgacccgacatcagtgacgtgttttggggagagcgcgggagccctcgtg	720
Db	820	TTCGGGGGTGACCCGACATCAGTGACGCTGTtTTGGGGAGAGCCGGGAGCCGCTCGGTG	879
Qy	721	ggcatgcacctgctgtccccccagccgggctgttccacagggccgtgctgcagagc	780
Db	880	GGCATGCACCTGTGTCCCCGCCACCGCGGGCCTGTtTCCACAGGGCCGTGCTGCAGAGC	939
Qy	781	ggtgcccccaatggaccctgggcccaggtgggcatgggagagggccctgcgagggccacg	840
Db	940	GGTGCCCCCAATGGACCCCTGGGCCACGGTGGCATGGGAGAGGCCCGTGCAGGGCCACG	999
Qy	841	cagctggcccacacctgtgggctgtcctccagggcgccactggtgggaatgacacagagctg	900
Db	1000	CAGCTGGCCACACTTGTGGGCTGTCTTCCAGGGCGCACTGGTGGGAATCACACAGAGCTG	1059
Qy	901	gtagcctgccttcggacacgacagcagcgaggtcctggtgaaccacgaatggcacgtgctg	960
Db	1060	GTAGCCTGCCTTCGGACACGACCGCAGCGCAGGTCTCTGTTGAACCAACGAATGGCACGTGCTG	1119
Qy	961	cctcaagaaagcgtcttccggttctccttctcgtgcctgtggtagatggagacttccctcagt	1020
Db	1120	CCTCAAGAAAGCGTCTTCCGGTTCCTCTCTGTCCTGTGCTGCTGTGGTAGATGGAGACTTCCCTCAGT	1179

Qy	1021	gacacccacagaggccctcatcaacgcgggagagacttccacggcctgcagtgctggtgggt	1080
Db	1180	GACACCCAGAGGCCCTCATCAACGGGAGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT	1239
Qy	1081	gtggtgaaggatgaggctcgtattttctggttttacggggcccccaggttcagcaaaagac	1140
Db	1240	GTGGTGAAGGATGAGGGCTCGTATTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAAGAC	1299
Qy	1141	aacgagtctctcatcagccgggcccagagttccttgcgggggtgcgggtcgggttccccag	1200
Db	1300	AACGAGTCTCTCATCAGCCGGGCCGAGTTCCTGGCCGGGTGCGGGTTCGCCAG	1359
Qy	1201	gtaagtgaacctggcagccgaggtctggttctctgcattacacagactggctgcaccccgag	1260
Db	1360	GTAAGTGACCTGGCAGCCGAGGCTGTGCTCTGCATTACACAGACTGGCTGCATCCCCGAG	1419
Qy	1261	gacccggcacgcctgagggagggccctgagcgtgtggtggcgacacaatgtcgtgtgc	1320
Db	1420	GACCCGGCAGCCTGAGGGAGGCCCTGAGCGATGTGTTGGCGCACCAATGTCTGTGTGC	1479
Qy	1321	cccgtggccacgtggtggcgactggtgcccaggtgcccgggtctacgcctacgtc	1380
Db	1480	CCCGTGGCCAGTGGTGGCGACTGGTGCCCCAGGGTGCCCCGGGTCTACGCTACGTC	1539
Qy	1381	tttgaacacggtgtcttccacgctctcttggccctctgtggatgggggtgccccacggctac	1440
Db	1540	TTTGAACACCGTGTCTCCACGCTCTCTCTGGCCCCCTGTGGATGGSGGTGCCCCACGGCTAC	1599
Qy	1441	gagatcgagttcatctttgggatccccctggacccctctcgaaactacacggcagaggag	1500
Db	1600	GAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG	1659
Qy	1501	aaaatcttcgcccagcgactgatgcatactgggccaaactttgcccgcacaggggatccc	1560
Db	1660	AAAATCTTCGCCACGACTGTATGCGATCTGGGCTTGGGCTTGGGCTTGGCGGATCCC	1719
Qy	1561	aatgagccccgagaccccaaggccccacaatggccccctacacggcggggctcagcag	1620
Db	1720	AATGAGCCCGAGACCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGCTCAGCAG	1779
Qy	1621	tacgttagcttgacctgcggccgctgaggtgcggggggctgcggccccagccctgc	1680
Db	1780	TACGTTAGTCTGGACCTGCGGCCGCTGAGGTGCGGGGGGCTGCGCGCCCGCCAGCCTGC	1839
Qy	1681	gccttctggaaccgcttctctcccccaaatgctcagcggccacctg	1724
Db	1840	GCCTTCTGGAACCGCTTCTCTCCCCAAATTGCTCAGCGCCACCCGG	1883

RESULT	3
HUMACHE	
LOCUS	HUMACHE 2218 bp mRNA linear PRI 09-JUN-1995
DEFINITION	Human acetylcholinesterase (ACHE) mRNA, complete cds.
ACCESSION	M55040
VERSION	M55040.1 GI:177974
KEYWORDS	acetylcholinesterase.
SOURCE	Human 21-week old fetus DNA, and cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 2218)
AUTHORS	Soreq,H.E., Ben-Aziz,R., Prody,C.A., Seidman,S., Gnatt,A., Neville,L., Lieman-Hurwitz,J., Lev-Lehman,E., Ginzberg,D., Lapidot-Lifson,Y. and Zakut,H.
TITLE	Molecular cloning and construction of the coding region for human acetylcholinesterase reveals a G + C-rich attenuating structure
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9688-9692 (1990)
MEDLINE	91088577
FEATURES	Location/Qualifiers
source	1..2218
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

mRNA	/map="7q22"			
	/dev_stage="21-week-old fetus"			
gene	/tissue_lib="lambda-gt10"			
	1. .2218			
CDS	/gene="ACHE"			
	/note="G00-118-746"			
sig_peptide	/product="acetylcholinesterase"			
	1. .2218			
mat_peptide	/gene="ACHE"			
	/EC_number="3.1.1.7"			
BASE COUNT	/codon_start=1			
	/product="acetylcholinesterase"			
ORIGIN	/protein_id="AAA68151.1"			
	/db_xref="GI:177975"			
Query Match	/db_xref="GDB:G00-118-746"			
	/translation="MRPQCLLHTPSLASPLLLLLLLGGGVGAEGREDAELLYTVR			
Best Local Similarity	99.8%;	Score 1722;	DB 9;	Length 2218;
Matches 1722;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	atgaggccccgcagtgctgtgcacacgcgcttccctggttcctcccaactccttcctc	60		
157	ATGAGGCCCCCGAGTGTCTGTGCACACGCCTTCCCTGGCTTCCCACTCCTTCTCCTC	216		
61	ctcctctggtcctgggtgagagtggtgggctgagggccggaggtgcagagctgctg	120		
217	CTCCTCTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGGATGCGAGAGTGTCTG	276		
121	gtgacggtgctggggccggtgcggggcaattcgctgaagacccccggggccctgtc	180		
277	GTGACGGTGCCTGGGGCCGGCTGCGGGGCATTCGCTGAAGACCCCGGGGCCCTGTC	336		
181	tctgcttctgggcatccccctttgcggagccaccccatgggaccccgctcttctgcc	240		
337	TCTGCTTCTCTGGCATCCCCCTTTCGGGAGCCACCCATGGGACCCCGTCGCTTCTGCCA	396		
241	ccggagcccaagcagccttggtcaggggtggttagacgctacaaccttccagagtgtcgc	300		
397	CCGGAGCCCCACAGCCCTTGGTCAGGGGTGGTAGACCGCTACAACCTTCCAGAGTGTCTGC	456		
301	taccaatatgtggacacccctataccccaggttttgagggcaccgagatgtggaaccccaac	360		
457	TACCAATATGTGGACACCCCTATACCCAGGTTTTCAGGGCACCCGAGATGTGGAACCCCAAC	516		
361	cgtgagctgagcagagactgcctgtacctcaactgtagacaccataccccggcctaca	420		
517	CGTGAGCTGAGCAGGAGTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCCTACA	576		
421	tccccacccctgtcctcgtctggtatctatgggggtggttctctacagtggggcctcctcc	480		

Db	577	TCCCCAACCCCTGTCTCTGTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC	636
QY	481	ttggacgtgtacgatggccgcttcttggtacagggccgagagactgtgctggtgccatg	540
Db	637	TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGTTGCCATG	696
QY	541	aactaccgggtgggagcccttggcttccctggccctgcccgggagccgagagccccgggc	600
Db	697	AACtACCCGGTGGAGCCCTTTGGCTTCCtTGGCCCTGCCGGGAGCCGAGAGCCCCGGGC	756
QY	601	aatgtgggtcctcctggatcagaggctggccctgagtgagggtgcaggaagaacgtggcagcc	660
Db	757	AATGTGGGTCTCTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCC	816
QY	661	ttcgggggtgacccgacatcagtgacgtgttttggggagagcgcggagccgcctcggtg	720
Db	817	TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCCTCGGTG	876
QY	721	ggcatgcacctgtgtccccccagcccggttccacagggccgctgtccacagggcgtgcagagc	780
Db	877	GGCATGCACCTGCTGTCCCCCAGCCCGGCGCTGTTCACAGGGCCGCTGCTGCAGAGC	936
QY	781	ggtgcccccaatggaccctgggccacggtgggcatgggagagggcccgctcgagggccacg	840
Db	937	GGTCCCCCAATGACCCCTGGGCCACGGTGGSCATGGAGAGGCCCGTCCGAGGGCCACG	996
QY	841	cagctggccccacctgtggctgtcctccagggcgccactggtgggaatgacacagagctg	900
Db	997	CAGCTGGCCCACTTGTGGCTGTCTCTCCAGCGGCACCTGGTGGGAATGACACAGAGCTG	1056
QY	901	gtagcctgccttcgggacacgaccagcgaggtccctggtgaaccacgaatggcacgtgctg	960
Db	1057	GTAGCCTGCCCTTCGGACACGACCAGCGCAGGTCTCTGTTGAACCAACGAATGGCACGTGCTG	1116
QY	961	cctcaagaaagcgtcttccggttctccttccttcctgctggtgtagatggagacttccctcagt	1020
Db	1117	CCTCAAGAAAGCGTCTTCCGGTTCTCCTTCGTGCCTGTGGTAGATGGAGACTTCCTCAGT	1176
QY	1021	gacacccacagggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt	1080
Db	1177	GACACCCACAGAGGCCCTCATCAACGCGGGAGACTTCACGGCCTGCAGGTGCTGGTGGT	1236
QY	1081	gtggtgaaggatgagggtcgtattttctgtgttacggggggccccaggcttcagcaagac	1140
Db	1237	GTGGTGAAGGATGAGGGCTCGTATTTTCTGTGTTACGGGCCCCAGGCTTCAGCAAAGAC	1296
QY	1141	aacgagtctctcatcagccggggccgaggttctctgcccgggtgcgggttccccag	1200
Db	1297	AACGAGTCTCTCATCAGCCGGCCGAGTTCCTTGCCGGGGTGCGGGTTCGGGTTCCCAG	1356
QY	1201	gtaagtgcctggcagccgaggtgtggtcctgcattacacagactggtgcacatccccag	1260
Db	1357	GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTGCATTACACAGACTGGCTGCATCCCGAG	1416
QY	1261	gacccggcacgctgagggagggccctgagcgtgtggtgggacccaatgtcgtgtgc	1320
Db	1417	GACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGGTGGCGACCAATGTGTGTGC	1476
QY	1321	cccgtggcccaagctggctggggcactggctcccaggtgccccgggtctacgcctacgtc	1380
Db	1477	CCCCGGCCACGCTGGCTGGCGACTGGCTGCCCCAGGGTGCCCCGGTCTACGCCTACGTC	1536
QY	1381	tttgaacacccgtgcttccacgctctcctcctggccctgtggtatgggggtgccccacggctac	1440
Db	1537	TTTGAACACCGTGTTCACGCTCTCCTGGCCCTGTGGATGGGGGTGCCCCACGGCTAC	1596
QY	1441	gagatcgagttcatctttgggatacccccttgagccccctctcgaaactacacagggagag	1500
Db	1597	GAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTTCGAAACTACACGGCAGAGG	1656
QY	1501	aaaatcttcgcccagcgactgatcgatacttggggccaaacttggccgacaggggatccc	1560

Db 1657 AAAATCTTCGCCCCAGCGACTGATGCGATACTGGGCCAACTTTGGCCCCGACAGGGGATCCC 1716
QY 1561 aatgagccccgagagacccccaaaggccccacaaatggcccccgctacacggcggggctcagcag 1620
Db 1717 AATGAGCCCCGAGACCCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1776
QY 1621 tacgttagctggacctgcggcgctggagtgcgggcggggctgcgccccaggcctgc 1680
Db 1777 TACGTTAGTCTGGACCTGCGGCGCTGGAGGTGCGGCGGGGCTGCGGCCCCAGGSCCTGC 1836
QY 1681 gccttctggaaccgcttcctcccaaatgtctcagcgccacc 1722
Db 1837 GCCTTCTGGAACCGCTTCCTTCCCCAAATTGCTCAGCGGCCACC 1878

RESULT 4
AR070205 AR070205 2256 bp DNA linear PAT 18-FEB-2000
LOCUS
DEFINITION Sequence 5 from patent US 5891725.
ACCESSION AR070205
VERSION AR070205.1 GI:7221093
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2256)
AUTHORS Soreq,H., Zakut,H. and Eckstein,F.
TITLE Synthetic antisense oligodeoxynucleotides and pharmaceutical compositions containing them
JOURNAL Patent: US 5891725-A 5 06-APR-1999;
FEATURES Location/Qualifiers
source 1. .2256
/organism="unknown"
BASE COUNT 390 a 757 c 680 g 429 t
ORIGIN

Query Match 99.8%; Score 1722; DB 6; Length 2256;
Best Local Similarity 100.0%; Pred. No. 3.4e-294;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgaggccccgcagtgctgtgcacacgccttcctccctggcttccccactccttcctc 60
Db 160 ATGAGCCCCCGCAGTGTCTGTGCACACGCCCTTCCTGGCTTCCCCACTCCTTCTCCTC 219
QY 61 ctctctggctcctgggtgagagtggtgggctgagggcgaggatgcagagctgtc 120
Db 220 CTCCTCTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCGGAGGATGCAGAGCTGCTG 279
QY 121 gtgaggtgtcgtggggcgcggtgcggggcattcgccctgaagaccccccggggccccgtgc 180
Db 280 GTGAGGTGCTGGGGCGCGCTGCGGGGCATTTCGCCCTGAAGACCCCCCGGGGCCCTGTC 339
QY 181 tctgcttctcctgggcatccccctttgcggagccaccctatgggacccccgtcgctttctgcc 240
Db 340 TCTGCTTCTCTGGGCATCCCCCTTTGCGGAGCCACCCCATGSGACCCCCGTCGCTTCTGCCA 399
QY 241 ccggagcccaagcagcccttggtcagggggtggtagacgtactacaacctccagagtgtctgc 300
Db 400 CCGGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTTACAACCTTCCAGAGTGTCTGC 459
QY 301 tacaatatgtggacaccctatataccagggttttgagggcaccgagatgtggaacccccaac 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcaggactgctgtacctcaacgtgtggacaccataccccggcctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCGGCCCTACA 579
QY 421 tccccacccccctgtcctcgtctggatctatgggggtggcttctctacagtggggctcctcc 480
Db 580 TCCCCACCCCCTGTCCCTCGTCTGSATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTTCC 639

QY 481 ttggacgtgtacgatggccgcttcttggttacaggccgagagggactgtgtggtccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGCCGAGAGGACTGTGTGTGTCCATG 699
QY 541 aactacgggtggagccctttggcttctcctggccctgcgggggagccgagagccccgggc 600
Db 700 AACTACCGGTGGGAGCCTTTTGGCTTCTTGGCCCTGCGCGGGAGCCGAGAGGCCCGGGC 759
QY 601 aatgtgggtctcctggatcagagctggccctgcagtgagtggtgcaggagaaacgtgagacc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttcgggggtgacccgacatcagtgacgctgtttgggagagcgcgaggagccctcggtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGAGCCGCTCGGTG 879
QY 721 ggcataccactgtctgtccccccagcggggctgttccacagggccgtgtgcagagc 780
Db 880 GGCAATGACCTGTCTCCCGGCCAGCGGGGCCCTGTTCACAGGGCCGTGTGCAGAGC 939
QY 781 ggtgcccccaatggacctgggcccaggtggcatgggagagggccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGGCCGCTGCAGGGCCACG 999
QY 841 cagctggccccaccttgtgggctgtcctccaggcggcactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCACCTTGTGGGCTGTCTCCAGCGGCACCTGGTGGGAATGACACAGAGCTG 1059
QY 901 gtgacctgcttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
Db 1060 GTAGCCTGCCCTTCGGACACGACCGACGACGAGGTCTCTGGTGAACACGAATGGCAGCTGCTG 1119
QY 961 cctcaagaaagcgtcttccggttctccttctcgtgcctgtggtagatggagacttcctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCTCCTTCTCGTGCCTGTGGTAGATGGAGACTTTCCTCAGT 1179
QY 1021 gacacccagagagccctcatcaacgggggagacttccacggcctgcaggtgtgtgggt 1080
Db 1180 GACACCCAGAGGCCCTCATCAACGCGGGAGACTTCCACGSCCTGCAGGTGTGTGGGT 1239
QY 1081 gtggtgaaggatgagggctcgatatatttctggtttacaggggccccaggcttcagcaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCCAGGCTTCAGCAAAGAC 1299
QY 1141 aacgagtctctcatcagcccgggcgagttcctgcggcggtgcggggtcggggttccccag 1200
Db 1300 AACGAGTCTCTCATCAGCCGGGGCCGAGTTCCTGGCCGGGTGCGGGGTTCGCCAG 1359
QY 1201 gtaagtacctggcagccgaggtgtgtcctgcattacacagactggctgcataccccag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCCAG 1419
QY 1261 gaccggcacgcctgagggagggcctgagcgatggtggggcgacacacaatgtcgtgtgc 1320
Db 1420 GACCCGACGCGCTGAGGGAGGCCCTGAGCGATGTGGTGGCGCACACAATGTCTGTGTGC 1479
QY 1321 cccgtggccccagctggctggcgactggctgcacaggggtgcgggggtctacgcctacgtc 1380
Db 1480 CCGGTGCCCCAGCTGGCTGGCGGACTGGCTGCCCAGGGTGCCTGGGTCTACGCCTACGTC 1539
QY 1381 tttgaacacccgtgcttccacgctctcctggccccctgtggatggggggtgccccacgggtac 1440
Db 1540 TTTGAACACCGGTGCTTCCACGCTCTCTGGCCCCCTGTGGATGGGGGTGCCCCACGGGTAC 1599
QY 1441 gagatcgagttcatctttgggataccccctggacccctctcgaaactacacggcagaggag 1500
Db 1600 GAGATCGAGTTTATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaatcttcgcccagcagactgatgagatactgggccaactttgcgcgcacaggggatccc 1560
Db 1660 AAAATCTTCGCCCCAGCGACTGATGGGATACTGGGCCAACTTTGGCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagagcccccaaggcccccaaatggcccccggtacacggcggggtcagcag 1620

```
Db 1720 AATGAGCCCGAGACCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
Qy 1621 tacgttagctggacctgcggccgctgagagtgcggggctgcccagggcctgc 1680
Db 1780 TACGTTAGTCTGGACCTGCGGCCGTGAGGCTGCGGGGGCTGCGGCCAGGCCTGC 1839
Qy 1681 gccttctggaaccgcttcctcccccattgctcagcgcacc 1722
Db 1840 GCCTTCTGGAACCGCTTCCTCCCCAAATGCTCAGCGCCACC 1881

RESULT 5
AR070207 AR070207 3016 bp DNA linear PAT 18-FEB-2000
LOCUS
DEFINITION Sequence 7 from patent US 5891725.
ACCESSION AR070207
VERSION AR070207.1 GI:7221095
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3016)
TITLES Soreq,H., Zakut,H. and Eckstein,F.
Synthetic antisense oligodeoxynucleotides and pharmaceutical
compositions containing them
JOURNAL Patent: US 5891725-A 7 06-APR-1999;
FEATURES Location/Qualifiers
source 1..3016
/organism="unknown"
BASE COUNT 497 a 1065 c 840 g 614 t
ORIGIN

Query Match 99.8%; Score 1722; DB 6; Length 3016;
Best Local Similarity 100.0%; Pred. No. 3.2e-294;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgagggcccccgcagtgctgctgcacagcgccttccttgcttcccccactccttcctc 60
Db 160 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCTTCCTGGCTTCCCACTCCTTCTCCTC 219

Qy 61 ctccctctggctcctgggtggaggagtgggggctgagggccgggagagatgcagagctgctg 120
Db 220 CTCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279

Qy 121 gtgacgtgctggggccgctgcggggcattcgcctgaagacccccggggccctgtc 180
Db 280 GTGACGGTGCCTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGCCCTGTC 339

Qy 181 tctgctttctgggcacatcccccttgcggagccaccccatgggaccccgctcgcttctgcca 240
Db 340 TCTGCTTCTCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTCTGCCA 399

Qy 241 ccggagcccaagcagccttggtcagggtggtagacgctacaaccttccagagtgtctgc 300
Db 400 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 459

Qy 301 taccaatatgtggaacaccctatacccaaggttttgaggggcaaccagatgtggaaccccaac 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAAACCCCAAC 519

Qy 361 cgtgagctgagcaggagactgcctgtacctcaacgtgtggacaccataccccggcctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTTGTGGACACCATAACCCCGGCCTACA 579

Qy 421 tccccacacctgtcctcgtctggtatctatgggggtggccttctacagtgggcctcctcc 480
Db 580 TCCCCACCCCTGTCTCGTCTGGATCATATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 639

Qy 481 ttggacgtgtacgatggccgcttcttggtacagggccgagaggactgtgctggttccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 699
```

```
Qy 541 aactaccgggtgggagccttggcttcttgccctgccgggggagccgagagccccgggc 600
Db 700 AACTACCGGGTGGGAGCCTTTGGCTTCTTGCCCTTCCGGGGAGCCGAGAGCCCCGGGC 759
Qy 601 aatgtgggtctcctggatcagaggctggccctgcagtgggtgcaggagaacgtggcagcc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGSCCCTGCAGTGGGTGCAGAGAACTGGCAGCC 819
Qy 661 ttcgggggtgaccccgacatcagtgacgtgttttggggagagcgcgggagccgcctcggtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 879
Qy 721 ggcattgcacctgctgtccccgcacagccgggggcctgttccacaggccgctgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCCGCCAGCCGGGGCCTGTTCACAGGCGCCTGCTGCAGAGC 939
Qy 781 ggtgcccccaatggacctgggcccacagtggggcatgggagagggccgctcgagggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGCTGGGCATGGGAGAGGCCGCTCGCAGGGCCACG 999
Qy 841 cagctggcccaccttggtggctgtcctcagggcggcactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCACCTTGTGGCTGTCTCCACGGCGGCACCTGGTGGGAATGACACAGAGCTG 1059
Qy 901 gtagcctgccttcggacacgaccagcaggtcctggtgaaccacgaatggcacgtgctg 960
Db 1060 GTAGCCTGCCTTCGGACACGACCAGCGCAGGTCTCTGGTGAACCCACCAATGGCAGCTG 1119
Qy 961 cctcaagaaagcgtcttccgggttctcctcgtgcctgtggtagatggagacttccctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCTCCTTCGTGCCTGTGGTAGATGGAGACTTCCTCAGT 1179
Qy 1021 gacacccccagaggccctcatcaacgcgggagacttccacggcctcaggtgctggtgggt 1080
Db 1180 GACACCCCAGAGGCCCTCATCAACGGCGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT 1239
Qy 1081 gtggtgaaggatgagggtcgtctatttctggtttacggggccccagccttcagcaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAAGAC 1299
Qy 1141 aacgagtctctcatcagccgggcccaggttccctggtgcgggggtgcgggtcgggttccccag 1200
Db 1300 AACGAGTCTCTCATCAGCCGGCGCGAGTCTCTGGCGGGGTGCGGGTTCGGGTTCGCCAG 1359
Qy 1201 gtaagtgcctggcagccgaggctgtgttcctgcattacacagactggctgcaccccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGCTCTGCATTACACAGACTGGCTGCATCCCGAG 1419
Qy 1261 gacccggcaacgctgagggagggccctgagcagatgtggtgggcgacacaaatgctgtgc 1320
Db 1420 GACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGGTGGCGGACCAACAATGTCTGTGTC 1479
Qy 1321 cccgtggccagctggctggcgactggctgcccagggtgcccgggtctacgcctacgtc 1380
Db 1480 CCCGTGGCCAGCTGGCTGGCGGACTGGCTGCCAGGGTGCCCGGGTCTACGCCTACGTC 1539
Qy 1381 tttgaacacggtgcttccacgctcctcctggccccctgtggatgggggtgccccacggctac 1440
Db 1540 TTGAACACCGTGTCTCCACGCTCTCTTGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599
Qy 1441 gagatcgagttcatctttgggatccccctggacccccctctcgaactacacagcgagaggag 1500
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCTCGAAACTACACGGCAGAGGAG 1659
Qy 1501 aaaatcttcgccccagcagctgatgcgatactgggccaactttgccgcgacaggggatccc 1560
Db 1660 AAAATCTTCGCCCCAGCGACTGATGCGGATACTGGGCCAACTTTTGCCCGCACAGGGGATCCC 1719
Qy 1561 aatgagccccgagagaccccaaggcccccaaatggccccccgtacacggggggtcagcag 1620
Db 1720 AATGAGCCCCGAGAGACCCCAAGGCCCCCAATAATGGCCCCCCCCGTACACGGCGGGGCTCAGCAG 1779
```

Qy 1621 tacgttagtctggacctggcgcgctggaggtgcggcgggggtgcgcgccaggcctgc 1680
|||||
Db 1780 TACGTAGTCTGGACCTGCGCGCTGGAGGTGCGCGGGGCTGCGGCCAGGCCTGC 1839

Qy 1681 gcctctggaaccgcttcctcccaaatgtctcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCCTCCCAAAATGCTCAGCGCCACC 1881

RESULT 6
AX275254
LOCUS AX275254 5767 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 3 from Patent WO0171014.
ACCESSION AX275254
VERSION AX275254.1 GI:16547674
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE 1 (sites)
AUTHORS Mor,T., Soreq,H., Arntzen,C. and Mason,H.
TITLE Expression of recombinant human acetylcholinesterase in transgenic plants

JOURNAL Patent: WO 0171014-A 3 27-SEP-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,
Tsafrir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,
Hugh S. (US)

FEATURES
source Location/Qualifiers
1..5767
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid vector pTM034"
BASE COUNT 1390 a 1521 c 1495 g 1361 t
ORIGIN

Query Match 99.6%; Score 1718.6; DB 6; Length 5767;
Best Local Similarity 99.8%; Pred. No. 1.1e-293;
Matches 1721; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 atgagggcccccgcagtgtctgtgcacacgccttcctggtccccactccttcctc 60
|||
Db 838 ATGGCTCCCCCGCAGTGTCTGTGCACACGCCTTCCCTGGCTTCCCCACTCCTTCTCCTC 897

Qy 61 ctccctcgtgcctcctgggtggagagtggggctgagggccgggaggtatgcagagctgtc 120
|||
Db 898 CTCCTGTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 957

Qy 121 gtgacggtgcgtggggccggctgcggggcattgcgctgaagacccccggggccctgtc 180
|||
Db 958 GTGACGGTGGTGGGGCCGGCTGGGGGCATTGCGCTGAAGACCCCCGGGGCCCTGTC 1017

Qy 181 tctgctttcctgggcaccccttatccaggttttgaggccaccctatgggacccccgcgtcttctgcc 240
|||||
Db 1018 TCTGCTTCTCTGGGCATCCCCCTTTCGGGAGCCACCCATGGGACCCCTCGCTTCTGCCA 1077

Qy 241 ccggagcccaagcagccttgctcaggggtggtagacgctacaaccttccagagtgtctgc 300
|||||
Db 1078 CCGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 1137

Qy 301 taccaatatgtggacacccctataccaggttttgaggccacgagatgtggaaccccaac 360
|||||
Db 1138 TACCAATATGTGGACACCCCTATACCCAGGTTTTCAGGGCCACCGAGATGTGGAACCCCAAC 1197

Qy 361 cgtgagctgacgaggactgcctgtacctcaacgtgtggacaccatacccccgacctaca 420
|||||
Db 1198 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCCTACA 1257

Qy 421 tccccaccctcgtcctcgtctggtatctatggggtggtcctctcacagtggggcctcctcc 480
|||||
Db 1258 TCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCCTCC 1317

Qy 481 ttggacgtgtacgatggccgcttcttggtacaggccgagaggactgtgctggtgccatg 540
|||||

Db 1318 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGCCGAGAGGACTGTGCTGGTGTCCATG 1377
|||||
Qy 541 aactaccgggtgggagcctttgttcctcctggcctgcgcctgcgggagccgagagccccgggc 600
|||||
Db 1378 AACTACCGGTTGGAGCCCTTGTGCTTCCTGGCCCTGCCGGGAGCCGAGAGCCCCCGGC 1437
|||||
Qy 601 aatgtgggtctcctggatcagaggtggccctgcagtggtgggtcagagagaacgtggcagcc 660
|||||
Db 1438 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCC 1497
|||||
Qy 661 ttcgggggtgacccgacatcagtgcacgtgttttgggagagcggggagccgcctcggtg 720
|||||
Db 1498 TTCGGGGGTGACCCGACATCAGTGCAGCTGTTTGGGAGAGCCGCGGAGCCGCTCGGTG 1557
|||||
Qy 721 ggcattgcacctgctgtccccgcagccggggcctgttccacaggggcccgtgctgcagagc 780
|||||
Db 1558 GGCATGCACCTGCTGTCCCCGCCACCGCGGGCCTGTTCACAGGGCCGTGCTGCAGAGC 1617
|||||
Qy 781 ggtgcccccaatggaccctgggcacggtggcatgggagagagcccgctgcagggccacg 840
|||||
Db 1618 GGTGCCCCCAATGGACCTTGGCCACAGGTGGGCATGGGAGAGGCCCGTCGCAGGGCCACG 1677
|||||
Qy 841 cagctggcccaacctgtgggtgtcctccagcggcactggtgggaatgacacagagctg 900
|||||
Db 1678 CAGCTGGCCACCTTGTGGGTGTCTCTCCAGCGGCACCTGTGGGAATGACACAGAGCTG 1737
|||||
Qy 901 gtagcctgccttcggacacgacagcgcgaggtcctggtgaaccacgaatggcacgtgctg 960
|||||
Db 1738 GTAGCCTGCCCTTCGGACACGACCCAGCGCAGGTCTCTGGTGAACACGAATGGCACCTGCTG 1797
|||||
Qy 961 cctcaagaagaagcgtcttcctcgttccttcctgcctgcctgtggtagatggagacttccctcagt 1020
|||||
Db 1798 CCTCAAGAAAGCGTCTTCCGGTTCCTCTCGTGCCTGTGGTAGATGGAGACTTCCTCAGT 1857
|||||
Qy 1021 gacacccagagggccctcatcaacgggggagagacttccacggcctgcaggtgctggtgggt 1080
|||||
Db 1858 GACACCCAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCTGACAGGTGCTGGTGGGT 1917
|||||
Qy 1081 gtggtgaaggatgagggctcgtatttctggtttacggggcccaaggcttcagcaaaagac 1140
|||||
Db 1918 GTGGTGAAGGATGAGGGCTCGTATTTCCTGGTTTACGGGGCCCAAGGCTTCAGCAAAAGAC 1977
|||||
Qy 1141 aacgagtctctcatcagccggcgaggtcctgcggcggtgcggggtgcgggttccccag 1200
|||||
Db 1978 AACGAGTCTCTCATCAGCCGGCGAGTTCTTGCCGGGTGCGGGTTCGGGTTCGCCAG 2037
|||||
Qy 1201 gtaagtgccttgccagccgaggtgtgttcctgcattacacagactggctgcattccccgag 1260
|||||
Db 2038 GTAAAGTGACCTGGCAGCCGAGGCTGTGGTCTTGCAATTACACAGACTGGCTGCATCCCCGAG 2097
|||||
Qy 1261 gaccggcacgcctgagggagggccctgagcgatgtggtggcgaccacaatgtcgtgtgc 1320
|||||
Db 2098 GACCCGGCACGCCCTGAGGGAGGCCCTGAGCGGATGTGGTGGGGCCACCAATAATGTCGTGTGC 2157
|||||
Qy 1321 cccgtggccagctggctggcgactggctgcccagggtgcccgggtctacgcctacgtc 1380
|||||
Db 2158 CCGGTGGCCAGCTGGCTGGCGCAGTGGCTGCCAGGGTGCCCGGGTCTACGCCTACGTC 2217
|||||
Qy 1381 ttgaaacacggtgtctccacgctcctcctggaccctctcgaaactacacgagagag 1440
|||||
Db 2218 TTTGAACACCGGTGCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGTGGCCCCACGGCTAC 2277
|||||
Qy 1441 gagatcaggttcatcttgggatccccctggaccctctcgaaactacacgagagag 1500
|||||
Db 2278 GAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTTACACGGCAGAGGAG 2337
|||||
Qy 1501 aaaatcttcgcccagcgactgatgcatactgggccaactttgcccgcacaggggatccc 1560
|||||
Db 2338 AAAATCTTCGCCCAGCGACTGATGCGATACTGGGCCAACTTTGGCCGCACAGGGGATCCC 2397
|||||
Qy 1561 aatgagcccccgagacccccaaaggccccacaatggccccctacacggcgggggtcagcag 1620
|||||

Db 2398 AATGAGCCCCGAGACCCCAAGGCCCCACAATGGCCCCCGTACACGGCGGGGCTCAGCAG 2457

QY 1621 tacgttagctggacctgcggccgctgaggtgcggcggggctgcgcgcccaggcctgc 1680

Db 2458 TACGTTAGTCTGGACCTGCGGCCGCTGAGAGGTGCGCGGGGCTGCGGCCACAGCCTGC 2517

QY 1681 gccttctggaaccgcttcctccccaattgctcagcgccacctga 1725

Db 2518 GCCTTCTGGAACCGCTTCCTCCCCAAATGCTCAGCGTACCTGA 2562

RESULT 7

AX275255

LOCUS AX275255 14446 bp DNA linear PAT 29-OCT-2001

DEFINITION Sequence 4 from Patent WO0171014.

ACCESSION AX275255

VERSION AX275255.1 GI:16547675

KEYWORDS .

SOURCE synthetic construct.

ORGANISM synthetic construct

artificial sequence.

REFERENCE 1 (sites)

AUTHORS Mor,T., Soreq,H., Arntzen,H., Arntzen,C. and Mason,H.

TITLE Expression of recombinant human acetylcholinesterase in transgenic plants

JOURNAL Patient: WO 0171014-A 4 27-SEP-2001;

BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor, Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason, Hugh S. (US)

FEATURES

source

1. .14446

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="plasmid vector pTM036"

BASE COUNT 3231 a 3831 c 4046 g 3042 t 296 others

ORIGIN

Query Match 99.6%; Score 1718.6; DB 6; Length 14446;

Best Local Similarity 99.8%; Pred. No. 9.2e-294;

Matches 1721; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 atgagcccccgagtgctgtcgtgcacagcgccttcctcggttccccactccttcctc 60

Db 142 ATGGCTCCCCCGCAGTGCTGTGTGCACAGCGCTTCCTGGCTTCCCACTCCTTCCTC 201

QY 61 ctccctcgtcctcctgggtggagagtggggctgagggccgggagatgcagagctgctg 120

Db 202 CTCCTCTGGCTCTGGGTGGAGAGTGGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 261

QY 121 gtgacggtgcgtggggcgcgctgcggggcatcgcctgaagacccccggggccctgtc 180

Db 262 GTGACGGTGCCTGGGGCGCGGCTTTCGGGAGCCACCCATGGGACCCCGGGGCCCTGTC 321

QY 181 tctgctttcctgggcatccccctttgcggagccacccatgggaccccgctcgtttctgcca 240

Db 322 TCTGCTTTCTTGGGCATCCCCCTTTTCGGGAGCCACCCATGGGACCCCGTCGCTTCTGCCA 381

QY 241 ccggagcccaagcagcccttggtcaggggtggtagacgctacaaccttccagagtgtcgc 300

Db 382 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 441

QY 301 taccaatatgtggacacccctataccagggttttgagggcaccgagatgtggaaccccaac 360

Db 442 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 501

QY 361 cgtgagctgagcaggagactgcctgtactcaacgtgtggaaccatacccccgccctaca 420

Db 502 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATAFACCCCCCGCCTACA 561

QY 421 tccccacacctgtcctcgtcgtggtatcagtggtggcttctacagtggggcctcctcc 480

Db 562 TCCCCCACCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 621

QY 481 ttggacgtgtacgatggccgcttcttggtacagggccgagaggactgtgctggtgccatg 540

Db 622 TTGGACGTGTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGGACTGTGTGTTCCATG 681

QY 541 aactaccgggtgggagccctttggcttccctggccctgcgggggagcgcgagagggccccgggc 600

Db 682 AACTACCGSGTGGGAGCCTTTGGCTTCTTGGCCCTGCCGGGAGCCGAGAGGCCCCCGGC 741

QY 601 aatgtgggtcctcctggatcagaggctggccctgcagtggtgctcagagaaacgtggcagcc 660

Db 742 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCC 801

QY 661 ttccgggggtgaccccgacatcagtgacgtctgttggggagagcgcggagccgcctcggtg 720

Db 802 TTCGGGGGTGACCCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGAGCCGCTCGGTG 861

QY 721 ggcatgcacctgctgtccccgcagcgcgggctgttccacagggcgctgctgcagagc 780

Db 862 GGCATGCACCTGTCTCCCGCCACGCCGGGCGCTGTTCCACAGGGCGGTGCTGCAGAGC 921

QY 781 ggtgcccccaatggacctgggcccacggtgggcatgggagggcccgctgcgagggccacg 840

Db 922 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGCATGGGAGAGGCCCTCGCAGGGCCACG 981

QY 841 cagctggccccacctgtgggctgtcctccagggcgccactggtgggaatgacacagagctg 900

Db 982 CAGCTGGCCCCACCTTGTGGCTGTCTCCAGCGGCACCTGGTGGGAATGACACAGAGCTG 1041

QY 901 gtacgctgccttcggacacgacagcagcaggtcctggtgaaaccacgaatggcacgtgctg 960

Db 1042 GTAGCCTGCCCTCGGACACGACACGCGCAGGTCCTGGTGAACCAACGAATGGCACGTGCTG 1101

QY 961 cctcaagaaagcgtcttcgggttctcctcgttcgtgctggtgtagatggagacttcctcagt 1020

Db 1102 CCTCAAGAAAGCGTCTTCCGGTTCTCCTTCGTCGCTGTGGTAGATGGAGACTTCCTCAGT 1161

QY 1021 gacacccccagagggccctcatcaacgcgggagagacttccacggcctgaggtgctggtgggt 1080

Db 1162 GACACCCCAAGAGGCCCTCATCAACGCGGAGAGACTTCCACGGCCTGCAGGTGCTGGTGGT 1221

QY 1081 gtggtgaaggatgagggctcgtctatttctggtttacggggccccaggttcagcaagac 1140

Db 1222 GTGGTGAAGGATGAGGGCTCGTATTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAAGAC 1281

QY 1141 aacgagtcctcctcatcagccggccgagttcctgcgggggtgcgggtcggggttccccacg 1200

Db 1282 AACGAGTCTTCATCAGCCGGCCCGAGTTCCTGGCCGGGTGCGGGTTCGGGTTCCCCAG 1341

QY 1201 gtaagtgccttgccagccgaggtggtgctcctgcattacacagactggctgcaccccgag 1260

Db 1342 GTAAGTGACCTGGCAGCCGAGGCTGTGTCTTCATACACAGACTGGCTGCATCCCGAG 1401

QY 1261 gacccggcaacgctgagggagccctgagcgtgctgcccaggtgcccgggtctacgcctacgtc 1320

Db 1402 GACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGGTGGCGACCAATGTCTGTGTGC 1461

QY 1321 cccgtggcccagctggctgggcgactggtgcccaggtgcccgggtctacgcctacgtc 1380

Db 1462 CCCGTGGCCAGCTGGCTGGCGGACTGGGTGCCAGGTTGCCGGGTCTACGCCTACGTC 1521

QY 1381 tttagaacacggtgcttccacgctctcctggccccctgtggatgggggtgccccacgggtac 1440

Db 1522 TTTGAACACCGTGCTTCCACGCTCTCCTGGCCCTGTGGATGGGGGTGCCCCACGGCTAC 1581

QY 1441 gagatcgcggttcatcttgggatccccctggacccccctctcgaaactacacgcgcagaggag 1500

Db 1582 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTTCGAAACTACACGGCAGAGGAG 1641

QY 1501 aaaaacttcgccccagcagctgatgcatactgggccaactttgcccgcacaggggatccc 1560

Db 1642 AAAATCTTCGCCACGCGACTGATGCGGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1701

Qy	1561	aatgagccccgagacccccaaagcccccacaaatgcccccggtacacggcggggctcagcag	1620
Db	1702	AATGAGCCCCGAGACCCCCAAGCCCCACAAATGGCCCCCGTACACGGCGGGGCTCAGCAG	1761
Qy	1621	tacgttagtctggacctgcggccgctggaggtgcggcgggggtgcgcgccaggcctgc	1680
Db	1762	TACGTAGTCTGGACCTGCGGCCGTGGAGGTGCGGGGGGCTGCGGCCCAGGCCTGC	1821
Qy	1681	gcctctggaaccgcttcctcccaaatgtctcagcgccacctga	1725
Db	1822	GCCTTCTGGAACCGCTTCCTCCCAAAATGCTCAGCGCTACCTGA	1866
RESULT 8			
OCU05036			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
polyA_signal			
polyA_site			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1465; Conservative			
Qy	91	gctgagggccgggaggatgcagagctgctggtgacggtgcgtggggggccgctgcggggc	150
Db	1	GCTGAGGSCGGGAGGACCCCGAGCTGCTGGTGACGGTTCCGGGGGGCCGGCTGCGCGGC	60

Qy	151	attgcctgaagacccccggggccctgtctctgtcttccctgggcatcccccttgcggag	210
Db	61	CTCCGTTTGAAGCCCCCGGGGGCCGGTCTCCGCTTCCTGGGATCCCTTCGAGGAG	120
Qy	211	ccaccatgggacccccgtcgtcttctgccaccggagcccaagcagccttggtcaggggtg	270
Db	121	CCACCCGTGGGCCCCCGTCGCTTCTTGCCCCCGGAGCCCAAGCGGCCCTGGCCGGGGTG	180
Qy	271	gtagacgtacaaacctccagagtctctgtctaccaataatgtggacacccctataccagggt	330
Db	181	CTGGACGCCACAGCCTTCCAGAGCGTCTGCTACCAATATGTGGACACCTGTACCCGGGC	240
Qy	331	tttgaggggcacccgagatgtggaacccaaccctcctctctgtgacgtgacgaggactgcctgtacctc	390
Db	241	TTTGAGGGCACTGAGATGTGGAACCCCAACCGTGAGCTGAGTGAGGACTGTCTGTACCTC	300
Qy	391	aacgtgtggacaccatacccccgccctacataccccccacctgtcctcgtctggtatctat	450
Db	301	AATGTATGGACACCATAACCCCGGCTACATCCCCACCCCTGTCTCGTCTGGATCTAC	360
Qy	451	gggggtgcttctacagtggggctcctcctctgtgacgtgtacgtgagcgccgcttcttggtta	510
Db	361	GGGGGTGGCTTCTACAGTGGGCCCTCCTCTTGACGCTGACTACGGCCGCTTCTCTGGTC	420
Qy	511	caggccgagagactgtgctggtgtccatgaactaccgggtgggagccttggcttctctg	570
Db	421	CAGGCCGAAGGACAGTGTCTGGTGGCCATGAACCTACCGGGTGGGAGCCTTTGGCTTCACG	480
Qy	571	gcccctgcgggggagccgagagggcccggggcaatgtgggtctctctggatcagaggctggcc	630
Db	481	TGCCTTCCGGGCAGCCGAGAGGCCCGCGGTAAATGTGGGCCCTGCTGGATCAGAGACTGGCG	540
Qy	631	ctgcagtgggtgcaggagaaactgtggagccttcgggggtgaccccgacatcagtgacgtg	690
Db	541	CTGCAGTGGGTGCAGGAGAAATGTGGCAGCCTTTGGAGGGGACCCAGCGTCAGTGACGCTG	600
Qy	691	tttggggagagcgcgggagccgctcgtgggtggcatgcacctgctgtcccccggccgagg	750
Db	601	TTTGGGAGAGCGCGGGCGCGCCTCAGTGGGCCCTGCACCTGCTGTCCCCGCCAGCCGG	660
Qy	751	ggcctgttccacagggcccgtgctgcagagcggtgcccccaatggaccctgggccacggtg	810
Db	661	GGCCTGTTCACACCGGCCGTGCTGCAGAGCGGGCGCCCCCAACGGACCCTGGGCCACAGTG	720
Qy	811	ggcatgggagagcccgctgcagggccacgcagctggccccacctgtggtggtgctgctcca	870
Db	721	GGCGTGGAGAGGCCCGCCGACGGGCCACGCTGCTGGCCCCCTGTGGTGTGTGCCCCCCA	780
Qy	871	ggcggcactggtgggaatgacacagagctgtagcctgccttcggacacgaccagcgag	930
Db	781	GGTGGCGCGGTGGCAACGACACAGAGTGGTGGCCTGCCTGCGGACCGCGCGGCTCAG	840
Qy	931	gtcctggtgaaccacgaatggcacgtgctgcctcaagaaagcgtcttcgggttctccttc	990
Db	841	GACCTGGTGGACCATGAGTGGCGTGTGCTGCCCCCAAGAAAGCATCTCCGCTTTTCTCTC	900
Qy	991	gtgcctggtgtagatggagacttctcctcagtgacacccccagagggccctcatcaacgcgga	1050
Db	901	GTGCCCCGTGGTAGACGGAGACTTCTCTAGTGACACGCCCGGCGGCTCATCAACGCCGA	960
Qy	1051	gacttccacggcctgcagggtgctggtgggtggtggaaggatgagggtcgtatttctg	1110
Db	961	GACTTCCAGGGCCTGCAGGTGCTGGTGGGTGTGGTGAAGGATGAGGGTACCTATTTTCTG	1020
Qy	1111	gtttacggggccccaggcttcagcaagacagaacagatctctctcatcagccgggcccagttc	1170
Db	1021	GTTTACGGGGCCCCGGGCTTCAGCAAAAGACAACGAATCTTTTCATAGCCGGGGCCAGTTC	1080
Qy	1171	ctggccgggggtgcgggttcgggttccccagggttaagtacacctggcagccgaggtgtggtc	1230
Db	1081	CTGGCTGGGTGCGGGTTCGGGGTCCCCCAGGCAAGCGACCTGGCGCGCGGCGGTGGTC	1140

```
QY 1231 ctgcattacagactggctgcatcccgaggaccggcacgccttgaggagccctgagc 1290
Db 1141 CTGCATTACAGACTGGCTGCACCCTGAGGACCCGGCCCGCTGAGGACGCCCTGAGC 1200

QY 1291 gatgtggtggcgaccacaatgtcgtgcccccgctgccccagctggctggcgactggct 1350
Db 1201 GACGTGGTGGCGACCACAAACGTGTCGCCCGCTGGCCCCAGCTGGCTGGCGACTGGCC 1260

QY 1351 gccagggtgcccgggtctacgcctacgtcttgaacacccgtgcttccacgctctcctgg 1410
Db 1261 GCCAGGGTGCCGGGTCTATGCCTACGCTCTTGAACACCCGGGCTTCCACGCTCTCCTGG 1320

QY 1411 cccctgtggtgggtgccccacggctacgagatcgagttcatctttgggatccccctg 1470
Db 1321 CCGCTCTGGATGGCGTGTCCCCACGGCTACGAGATCGAGTTCAGTTTCATCTTCGGGCTGCCCTTG 1380

QY 1471 gaccctctgaaactacacggcagagagaaaaatcttcgccccagcactgatcgatac 1530
Db 1381 GAACCCCTCGCTGAACCTACACCGAGGAGAGAGGATCTTTGCCCAGAGATTGATGAGATAC 1440

QY 1531 tgggccaactttgccccgcacaggggatcccaatgagccccgagaccccccaagccccacaa 1590
Db 1441 TGGGCCAACTTTGCACGCACACAGGGACCCGAACGAGCCCCCGCGACGCCAACGGCCCCGAG 1500

QY 1591 tggccccctacacgcgggggtcagcagtagcttagtctggaacctgcggccgctggag 1650
Db 1501 TGGCCACCGTACACGGCGGGGGCGCAGCAGTACGTTAGCCTGAACCTGCGGCCGCTGGAG 1560

QY 1651 gtgcggcgggggtgcgcgccaggcctgcgccttctggaaccgcttctccccaattg 1710
Db 1561 GTCGGAGGGGCTGCGCGCCCGCAGGCTTCTGGAACCGCTTCTGCCCCGCTGCTGCCAAACTG 1620

QY 1711 ctcagcgccacc 1722
Db 1621 CTCAGCGCCACC 1632

RESULT 9
AX306139
LOCUS AX306139 2089 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 890 from Patent WO0188188.
ACCESSION AX306139
VERSION AX306139.1 GI:17645432
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 890 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source
1..2089
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 409 a 638 c 596 g 446 t
ORIGIN
Query Match 74.3%; Score 1282; DB 6; Length 2089;
Best Local Similarity 84.0%; Pred. No. 1.4e-216;
Matches 1447; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 1 atgagggcccgcgagtgctgctgcacagcgcttccctggcttcccgactctctcctc 60
Db 1 ATGAGGCCCTCCCTGGTATCCCTGCATACACCTTCCCTGGCTTTCCACTCCTCTTCCTC 60

QY 61 ctcctctgctcctgggtggaggagtgggggtgagggccgggagagatgcagagctgctg 120
Db 61 CTCCTCTCCCTCCTGGGAGGAGGGGCAAGGGCTGAGGGCCGGGAAGACCCCGCAGCTGCT 120
```

```
QY 121 gtgacggtgcgtggggccggctgcggggcattcgctgaagacccccgggggcccctgtc 180
Db 121 GTGAGGGTTTCGAGGGGGCCAGCTGAGGGGCATCCGCCTGAAGGCCCTTGAGGCCCAGTC 180

QY 181 tctgctttctcgtgggcatacccccttttgcggagccacccccatgggaccccgctcgtcttctgtcca 240
Db 181 TCAGCTTTTCTGGGCATCCCTTTTCAGAGACCACCTGTGGGCTCACGTAGATTATGCCA 240

QY 241 ccggagcccaagcagccttggtcaggggtggttagacgctacaaaccttcacagagtgtctgc 300
Db 241 CCAGAGCCCAAGCGGCCCTGGTCAAGAGTGTGGATGCTACCACCTTCCAAAATGTCTGC 300

QY 301 taccaatatgtggacacccctataccccaggttttgaggccacccgagatgtggaaccccaac 360
Db 301 TACCAGTACGTGGACACCCTGTACCCCTGGTGTGAGGTACTGAGATGTGAACCCCAAC 360

QY 361 cgtgagctgagcaggaactgcctgtacctcaacgtgtggacaccataccccccggcctaca 420
Db 361 CGAGAGTTGAGTGAAGACTGCCCTGTATCTTAATGTGTGGACACCATACCCAGACCTGCT 420

QY 421 tccccacccccctgtcctcgtctggtgatctatgggggtggcttctacagtggggcccctctcc 480
Db 421 TCTCCACACCTGTCTCTCATCTGGATCTATGGGGGTGTTTCTACACGGAGCGGCCCTCC 480

QY 481 ttggacgtgtacgagtggccgcttcttgggtacagggccgagaggaactgtgctggtgtccatg 540
Db 481 TTGGATGTGTATGACGGCCGTTTCTTGCCCCAGGTTGAGGGAGCTGTGTGGTATCTATG 540

QY 541 aactaccgggtgggagccttggcttctcctggccccctgcggggagccgagagagccccgggc 600
Db 541 AACTACCGAGTGGGAACCTTTTGGCTTCTTGCCCTTACCAGGAAGCAGAGAAGCCCTTGGC 600

QY 601 aatgtgggtctcctcgtgatcagaggctggccccctgcagtggtggtgcagagaaactggtgcagcc 660
Db 601 AATGTAGTCTGCTGGATCAACGGCTTGCTTGCATTGCAATGGGTGCAAGAAAATATTGCAGCC 660

QY 661 ttcgggggtgacccccacatcagtgacgtgttctggggagagcgcgggagccgcccctcgggtg 720
Db 661 TTTGGGGGCGACCCGATGTCAGTGACTCTGTTTGGGGAGAGTGGGGTGCAGCCTCCGTG 720

QY 721 ggcattgcacctgctgtccccgcgccagccggggggttctccacagggccgctgctgcagagc 780
Db 721 GGCATGCACATACTGTCCCTGCCACGACGAGGAGCCCTCTTCCACAGGGCTGTCTCTCCAGAGT 780

QY 781 ggtgcccccaatggaccctgggccacggtggtgggcatgggagagggcccgctgcagggccacg 840
Db 781 GGCACACCCCAATGGGCCCTTGGCCACTGTGAGTGTGTGAGAGGCCAGCGCAGGGCCACA 840

QY 841 cagctggcccaactgtgtggctgtcctccagggcggtggtggtggaatgacacagagctg 900
Db 841 CTGCTGGCCCGCTTGTGGGCTGTCCCCCAGGTGGCGCTGTTGGCAATGACACCCGAGCTG 900

QY 901 gtagcctgccttcggacacgacgacgagcaggtcctggtggaaccacgaatggcacgtgctg 960
Db 901 ATAGCCTGCTTGAGGACAAGGCCCGCTCAGGACCTGGTGGACCACGAGTGGCACGTCTCTG 960

QY 961 cctcaagaaagcgtcttccgggttctccttcgtgcctgtggttagatggagacttccctcagt 1020
Db 961 CCTCAAGAAAGTATCTTCCGATTTTCCCTTCGTGCCTGTGGTAGACGGGGACTTCCCTCAGT 1020

QY 1021 gacacccccagaggccctcatcaaacgcgggagacattccacggcctgcagggctgctggtgggt 1080
Db 1021 GACACACCGGAGGCTCTCATCAATACTGAGATTTTCAAGACCTGCGAGGTGCTGGTGGGT 1080

QY 1081 gtggtgaagatgagggtcgtatttcttggtttacggggccccagggcttcagcaaaagac 1140
Db 1081 GTGGTGAAGGACGAGGGCTCCTACTTTCTGGTTTACGGGGTCCCAGGCTTCAGCAAAAGAC 1140

QY 1141 aacgagtctctcatcagccggcgagttcctggcgggtggtgcgggtcggtgtccccag 1200
Db 1141 AATGAATCTCTCATCAGCCGGGCCAGTTCCTGGCTGGGGTGGGATCGGTGTACCCCAA 1200

QY 1201 gtaagtgaacctgggcagccgaggctgtggtcctgcattacacagagactggtgcaccccgag 1260
```


Db 721 GGCATGCACATACTGTCCCTGCCACGACGAGACCTCTTCCACAGGGCTGTCTCCTCCAGAGT 780

QY 781 ggtgcccccaatggacccttggtggccacggtgggcatgggagagggccctgcgagggccacg 840

Db 781 GGCACACCCCAATGGCCCTTGGCCACTGTGAGTGCTGGAGAGGCCGCGCAGGGCCACA 840

QY 841 cagctggccccaccttggtggctgtcctccaggcgccactggtgggaatgacacagagctg 900

Db 841 CTGCTGGCCCGCCTGTGGGCTGTCCCCAGGTGGCGCTGGTGGCAATGACACCGAGCTG 900

QY 901 gttagcctgccttcggacacgacgacgagcgaggttcctggtgaaccacgcaatggcacgtgctg 960

Db 901 ATAGCCTGCTTGAGGACAAAGCCCGCTCAGGACCTGGTGGACCACGAGTGGCACGTCCTG 960

QY 961 cctcaagaagagcttccggttctccttcctgcctgtggtagatggagacttcctcagt 1020

Db 961 CCTCAAGAAACTATCTTCCGATTTTCTCGTGTGCTGTGGTAGACGGGACTTCCTCAGT 1020

QY 1021 gacacccccagggccctcatcaacgcgggagagacttccacggcctgcaggtgctggtgggt 1080

Db 1021 GACACACCGGAGGCTCTCATCAATACTATGAGATTTTCAAGACCTGCAGTGCTGTGGGT 1080

QY 1081 gtggtgaagatgagggctcgtattttctggttttacggggccccaggcttcagcaaaagac 1140

Db 1081 GTGGTGAAGGACGAGGGCTCCTACTTTCTGGTTTACGGGGTCCCAGGCTTCAGCAAGAC 1140

QY 1141 aacgagttctctcatcagccgggcccaggttctcctggccgggtgcgggttcctccacg 1200

Db 1141 AATGAATCTCTCATCAGCCGGGCCAGTTCTCTGGCTGGGTGCGGATCGGTGTACCCCAA 1200

QY 1201 gtaagtacctggcagccgaggtgtgttctctgcattacacagactggctgcacccccgag 1260

Db 1201 GCAAGTGACCTGGCGGCCGAGGCTGTGTCTGCATTACACAGACTGGTGTGCACCCCTGAG 1260

QY 1261 gacccggcacgcctgagggagggccctgagcgatggtggggcgacacacaatgtcgtgtgc 1320

Db 1261 GACCCTACTACCTGAGAGATGCCATGATGTCAGTGGTAGGCGACCAACAACGTTGTGTGC 1320

QY 1321 cccgtggccccagctggctgggcgactggctgccccaggggtgccccgggtctacgcctacgtc 1380

Db 1321 CCTGTGGCCCCAGCTGGCTGGCGACTGGCTGCCCAAGGGCCCCGGGTCTATGCCTACATC 1380

QY 1381 tttgaacaccgtgcttccacgctctctcctggccccctgtggatgggggtgccccacggctac 1440

Db 1381 TTTGAACACCGTGCCTCCACACTGACTTGGCCCCCTCTGGATGGGGTGCCCCCATGGCTAT 1440

QY 1441 gagatcgagttcatcttttgggatccccctggaccctctctcgaaaactacacggcgagaggag 1500

Db 1441 GAAATCGAGTTCACTCTTTGGGCTCCCCCTCGATCCCTCGCTGAACTACACACGGAGGAG 1500

QY 1501 aaaatcttcgcccagcactgatgcgatactgggccaaactttgccccgacaggggatccc 1560

Db 1501 AGGATCTTTGCTCAGCGACTTATGAATACTGGACCAATTTTGCCCGCACAGGGGACCCC 1560

QY 1561 aatgagccccgagaccccccaaggcccccaaatggcccccggtacacggcgggggtcagcag 1620

Db 1561 AATGACCCCTCGAGACTCCAATACTCCACAGTGGCCACCGTACACCACTGCCGCGCAGCAA 1620

QY 1621 tacgttagtctggacctgcggccgctgagaggtgcggcggggctgcgcgccccagcctgc 1680

Db 1621 TATGTGAGCCCTGAACCTGAAGCCCTTAGAGGTGCGGGGGGACTGCGGCCCCCAGACCTGC 1680

QY 1681 gccttctggaaccgcttctcccccaaatgtctcagcgccacc 1722

Db 1681 GCCTTCTGGAATCGCTTTCTCCCCAAATTGCTCAGCGGCCACC 1722

RESULT 11

S50879

LOCUS S50879 2066 bp mRNA linear ROD 08-MAY-1993

DEFINITION acetylcholinesterase T subunit [rats, mRNA Partial, 2066 nt].

ACCESSION S50879

VERSION S50879.1 GI:262092

KEYWORDS Rattus sp.

SOURCE Rattus sp.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 2066)

AUTHORS Legay,C., Bon,S., Vernier,P., Coussen,F. and Massoullie,J.

TITLE Cloning and expression of a rat acetylcholinesterase subunit: generation of multiple molecular forms and complementarity with a Torpedo collagenic subunit

JOURNAL J. Neurochem. 60 (1), 337-346 (1993)

MEDLINE 93107932

REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 121005] from the original journal article. This sequence comes from Fig. 1.

FEATURES

Location/Qualifiers

1..2066

/organism="Rattus sp."

/db_xref="taxon:10118"

1..2066

/gene="acetylcholinesterase T subunit, AChE"

1..1845

/gene="acetylcholinesterase T subunit, AChE"

/note="This sequence comes from Fig. 1; AChE"

/codon_start=1

/product="acetylcholinesterase T subunit"

/protein_id="AAB24586.1"

/db_xref="GI:262093"

/translation="MRPPWYPLHTPSLASPLFLLLSLGGGARAEGREDPQLLVVRV

GGQLRGIRLKAPGVPVSAFLGIPFAEPPVGSRRFMPPEKRPKWSGILDATTFQNVCYQ

YVDTLYPGFEGTEMNPNRELSDECLYNVWTPYPRPTSPTPVLWIYGGFYSGASS

LDVYDGRFLAQVEGTVLVSMNYRVGTGFLALPGSREAPGNVGLDQRLALQWQENI

AAFGGDPMSVTLFGESAGASVGMHILSLPSRSLFHRVLOSCTPNGPWATVSAGEAR

RRATLLARLVGCPGGAGGNDTELISCLTRPAQDLVDHEWHVLPQESIFRFSFVPV

DGDFLSDTPDALINTGDFQLQVLGVVKDEGSYFLVGVPGFSKDNESLISRAQFLA

GVRIGVPOASDLAAEAVVLHVTDWLHPEDPAHLRDMASAVVDHNVCPVAQLAGRLA

AQARVYAYIFEHRASLTLPWMGMVPHGYEIEFIFGLDLPDSLNYTVEERIFAQRLM

OQWTFARTGDPNDRDSKSPWPYYTTAAQQYVSLNKLPLEVRRGLRAQTCAFWNR

LPKLLSATDTLDEAERQWKAEEFHRWSSYMVHWKNQFDHYSKQERCSDL"

BASE COUNT 394 a 629 c 590 g 453 t

ORIGIN

Query Match 74.0%; Score 1276.2; DB 10; Length 2066;

Best Local Similarity 83.8%; Pred. No. 1.4e-215;

Matches 1443; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctctgtgcacacgccttccttcctgcttcctccactccttcctc 60

Db 1 ATGAGGCCTCCTCGTATCCCTGCATACACCTCCCTCGCTTCTCCACTCCTCTTCCTC 60

QY 61 ctcctctggctcctgggtgagagagtggggctgagggcgccggaggatgcagagctgctg 120

Db 61 CTCCTCTCCCTCCTGGGAGGAGGGGCAAGGGCTGAGGGCCGGGAAGACCCCTCAGCTG 120

QY 121 gtgacggtgcgtggggggccggctgcggggcattcgctgaagacccccggggggccctgtc 180

Db 121 GTGAGGGTTCGAGGGGGCCAGCTGAGGGGGCATCCGCCCTGAAGCCCCCTGGAGCCC 180

QY 181 tctgcttctcctgggcatcccccttttgcggagggcaccacctgggacccccgtcgtcttc 240

Db 181 TCAGCTTTTCTGGGCATCCCCCTTTGCAGAGAGCCACCTGTGGGCTCACGTAGATTAT 240

QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttccagagtgctctgc 300

Db 241 CCAGAGCCCCAAGCCCCCTTGGTCAGGAATATTGGATGCTACCACCTTCCAAATGTCT 300

QY 301 taccaatatgtgacacccctataccagggttttgaggggcaccgagatgtggaaccccaac 360

Db 301 TACCAATACGTGGACACCCCTGTACCCCTGGGTTTGAGGGTACCAGAGATGTGGAACCC 360

QY 361 cgtgagctgagcagaggactgcctgtacctcaactgcgtggacacccataccccggcctaca 420

Feature	Sequence	Position	Count	Base Count	Origin
intron	1994. .2988 /gene="ACHE" /note="G00-118-746" /number=3 /evidence=experimental 2309 /gene="ACHE" /note="g or a; within Alu repeat; g more common; polymorphism at 2309" /replace="a" 2989. .3158 /gene="ACHE" /note="G00-118-746" /number=4 /evidence=experimental 3159. .3238 /gene="ACHE" /note="G00-118-746" /number=4 /evidence=experimental 3239. .3991 /gene="ACHE" /note="includes retained intron 5; G00-118-746" /number=5 /evidence=experimental 3290 /gene="ACHE" /note="Pro 561 more common; C/G polymorphism at Pro561(CCG) or Arg 561(CGG)" /replace="g" 3992. .>4185 /gene="ACHE" /note="G00-118-746" /number=6 /evidence=experimental	792 a	1320 c	1236 g	837 t
variation					
exon					
intron					
exon					
variation					
exon					

Query Match 69.4%; Score 1198; DB 9; Length 4185;
Best Local Similarity 81.8%; Pred. NO. 7.8e-202;
Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;

QY	1	atgagggccccgcagtgctgtgcacacgaccttccctggcttcccaactccttctcctc	60
DB	95	ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCACCTCTTCTCCTC	154
QY	61	ctcctctggctcctgggtgagggagtgggggctgagggccgggagggatgcagagctgctg	120
DB	155	CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG	214
QY	121	gtgacggtgcgtggggccggctgcggggcattgcctgaagacccccggggccctgtc	180
DB	215	GTGACGGTGC GTGGGGCCGGCTGCGGGGCATTTCGCCTGAAGACCCCCGGGGCCCTGTC	274
QY	181	tctgcttctgggcatacccccttgcggagccacccatgggaccgcgtcttcttgcca	240
DB	275	TCTGCTTTCTGGGCATCCCCCTTTCGGGAGCCACCCATGGGACCCGTCGCTTTC TGCCA	334
QY	241	ccggagcccaagcagccttggtcagggggtggtagacgcctacaaccttccagagtgtctgc	300
DB	335	CCGGAGCCCAAGCAGCCTTGTGTAGGGGTGTAGACGCTACACCTTCCAGAGTGTCTGC	394
QY	301	taccaatatgtggacacccctataccccagggtttttaggggcacccagatgtggaaccccaac	360
DB	395	TACCAATATGTGGACACCCCTATACCCAGGTTTTTGAGGGCACCGAGATGTGGNAACCCCAAC	454
QY	361	cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccgccctaca	420
DB	455	CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGCTGTGGACACCATACCCCGGCCTACA	514
QY	421	ccccccacccctgtcctcgtctggatctatgggggtggcttctacagtggggcctcctcc	480

```
Qy 1215 agcgaaggctgtggtcctgcattacacagactggctgcatccgagggacccggcacgcct 1274
|||||
Db 1655 AGCGAGGCTGTGGTCTGCTATACACAGACTGGCTGCATCCGAGGACCCGGCACGCCT 1714
|||||
Qy 1275 gaggagggccctgagcgatgtgtgggcgaccacaatgtcgtgtccccgtggcccagct 1334
|||||
Db 1715 GAGGAGGCCCTGAGCGATGTGGTGGCGGACCACACAATGTGCTGCCCCCGTGGCCAGCT 1774
|||||
Qy 1335 ggctgggcgactggctgcccaggggtgccccgggtctacgcctacgtctttgaacaccgtgc 1394
|||||
Db 1775 GGCTGGCGGACTGGCTGCCAGGGTGGCCGGGTCTACGCCTACGTCTTTGAACACCCGTGC 1834
|||||
Qy 1395 ttccagctctcctggccccctgtgatgggggtgccccacggctacgagatcgagttcat 1454
|||||
Db 1835 TTCCACGCTCTCCTGGCCCCCTGTGATGGGGGTGCCCCACGGCTACGAGATCGAGTTTCAT 1894
|||||
Qy 1455 ctttgggattccccctggaccctctcgaactacacggcgagaggagaaaatcttcgccc 1514
|||||
Db 1895 CTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAGAAATCTTCGCCCA 1954
|||||
Qy 1515 gcgactgatcgatactggggccaaactttgcccgcacaggg 1554
|||||
Db 1955 GCGACTGATCGGATACTGGGCCAACTTTGCCCGCACAGGG 1994
|||||

RESULT 13
AC011895/c
LOCUS AC011895 172358 bp DNA linear PRI 30-SEP-2000
DEFINITION Homo sapiens BAC clone RP11-126L15 from 7q22, complete sequence.
ACCESSION AC011895
VERSION AC011895.4 GI:9454624
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 172358)
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
2 (bases 1 to 172358)
Sun,H., Nguyen,C., Drone,K., Laplant,Y. and Reitz,L.
The sequence of Homo sapiens BAC clone RP11-126L15
Unpublished
3 (bases 1 to 172358)
Waterston,R.H.
Direct Submission
Submitted (15-OCT-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 172358)
Waterston,R.H.
Direct Submission
Submitted (26-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 172358)
Waterston,R.H.
Direct Submission
Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 172358)
Waterston,R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 26, 2000 this sequence version replaced gi:7630854.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
```

Web site: <http://genome.wustl.edu/gsc>
Contact: saplensewatson.wustl.edu
----- Summary Statistics
Center project name: H_NH0126L15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/OTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenio,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>) VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-336D7. Actual start of this clone is at base position 1 of RP11-126L15; actual end is at base position 172358 of RP11-126L15.

The clone RP11-126L15 contains a transposon that inserts at base position 30153, which has been omitted from the submitted sequence.

There are polymorphic base pair differences in the overlap between the clone RP11-126L15 and RP11-336D7.

FEATURES

source

Location/Qualifiers
1. .172358
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q22"
/clone="RP11-126L15"
/clone_lib="RPCI-11"

repeat_region

1. .249

/rpt_family="Alu"

repeat_region

935. .1075

/rpt_family="Alu"

repeat_region

1076. .1390

/rpt_family="Alu"

repeat_region

1391. .1553

/rpt_family="Alu"

repeat_region

2087. .2371

/rpt_family="Alu"

repeat_region

3027. .3324

/rpt_family="Alu"

repeat_region

3357. .3491

/rpt_family="Alu"

repeat_region

3492. .3795

/rpt_family="Alu"

```
repeat_region 3799. .4102
repeat_region /rpt_family="Alu"
repeat_region 4550. .4853
repeat_region /rpt_family="Alu"
repeat_region 4879. .5089
repeat_region /rpt_family="Alu"
repeat_region 5107. .5403
repeat_region /rpt_family="Alu"
repeat_region 5495. .5782
repeat_region /rpt_family="Alu"
repeat_region 6357. .6656
repeat_region /rpt_family="Alu"
repeat_region 6661. .6848
repeat_region /rpt_family="Alu"
repeat_region 6849. .7025
repeat_region /rpt_family="Alu"
repeat_region 7770. .8089
repeat_region /rpt_family="Alu"
repeat_region 8110. .8419
repeat_region /rpt_family="Alu"
misc_feature 8185. .8623
misc_feature /note="similar to
repeat_region 9602. .9902
repeat_region /rpt_family="Alu"
repeat_region 10324. .10630
repeat_region /rpt_family="Alu"
repeat_region 11217. .11526
repeat_region /rpt_family="Alu"
repeat_region 11667. .11976
repeat_region /rpt_family="Alu"
repeat_region 12357. .12655
repeat_region /rpt_family="Alu"
repeat_region 12657. .12781
repeat_region /rpt_family="Alu"
repeat_region 12783. .12955
repeat_region /rpt_family="Alu"
repeat_region 12956. .13236
repeat_region /rpt_family="Alu"
repeat_region 13237. .13367
repeat_region /rpt_family="Alu"
repeat_region 13403. .13444
repeat_region /rpt_family="Alu"
repeat_region 13445. .13547
repeat_region /rpt_family="Alu"
repeat_region 13561. .13864
repeat_region /rpt_family="Alu"
repeat_region 14215. .14326
repeat_region /rpt_family="Alu"
repeat_region 14327. .14591
repeat_region /rpt_family="Alu"
repeat_region 14592. .14674
repeat_region /rpt_family="L1"
repeat_region 14675. .14966
repeat_region /rpt_family="Alu"
repeat_region 14967. .15001
repeat_region /rpt_family="L1"
repeat_region 15010. .15337
repeat_region /rpt_family="Alu"
repeat_region 16240. .16509
repeat_region /rpt_family="Alu"
repeat_region 16522. .16656
repeat_region /rpt_family="Alu"
repeat_region 16687. .16978
repeat_region /rpt_family="Alu"
repeat_region 17053. .17344
repeat_region /rpt_family="Alu"
repeat_region 18167. .18444
repeat_region /rpt_family="Alu"
repeat_region 18848. .19163
repeat_region /rpt_family="Alu"
misc_feature 19101. .19232
misc_feature /note="similar to
repeat_region 19172. .19471
```

EST AI539956 (NID:g4457329)"

EST BE160273 (NID:g8622994)"

```
repeat_region /rpt_family="Alu"
repeat_region 19482. .19793
repeat_region /rpt_family="Alu"
repeat_region 19889. .20004
repeat_region /rpt_family="Alu"
repeat_region 20155. .20444
repeat_region /rpt_family="Alu"
repeat_region 20455. .20725
repeat_region /rpt_family="Alu"

Query Match 69.4%; Score 1198; DB 9; Length 172358;
Best Local Similarity 81.8%; Pred. No. 3.7e-202;
Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;

QY 1 atgagggcccgagtgctgtgcacacagccttccctgctgcttccccactccttcctc 60
|
Db 136364 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCTTCCCTGGCTTCCCCACTCCTTCTCCTC 136305

QY 61 ctccctctggtcctcctggtgaggagtggtgggctgagggccggaggagtcagagctgctg 120
|
Db 136304 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 136245

QY 121 gtgacggtgcgtggtggggccggctgcgggctgcgggcattgcctgaagacccccgggggccccctgtc 180
|
Db 136244 GTGACGGTGCCTGGGGGCCGGCTGCGGGGCAATTCGCCTGAAGACCCCCGGGGGCCCTGTC 136185

QY 181 tctgcttctcctgggcattccccctttgcggagccaccccatgggacccccgtcgtcttctgccca 240
|
Db 136184 TCTGCTTTCTCTGGGCATCCCCTTTGCGGAGCCACCCCATGGGACCCCGTCTGCTTCTGCCCA 136125

QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaacctccagagtgctctgc 300
|
Db 136124 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGCTAGACGCTACAACCTTCCAGAGTGTCTGC 136065

QY 301 taccaatatgtggacacccctataccccaggttttgaggccacccagagatgtggaaccccccaac 360
|
Db 136064 TACCAATATGTGGACACCCCTATACCCAGSTTTTGAGGGCACCGAGATGTGGAACCCCAAC 136005

QY 361 cgtgagctgagcggaggactgcctgtacctcaacgctgtggacaccataccccccggccctaca 420
|
Db 136004 CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGCTGTGGACACCATACCCCGGCCCTACA 135945

QY 421 tccccacccctgtcctcgtctggtgatctatggtgggtggtggtcttctacagtggtggccctctcc 480
|
Db 135944 TCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCCTCCTCC 135885

QY 481 ttggacgtgtacgatggccgctcttggtaaggccgagaggaactgtgctggtgccatg 540
|
Db 135884 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGTGGTGTCCATG 135825

QY 541 aactaccgggtgggagcctttggcttctcctggccctgcgggggagccgagagggccccgggc 600
|
Db 135824 AACTACCGGTTGGGAGCCTTTGGCTTCTCTGGCCTTCCCGGGAGCCGAGAGGCCCGCGGC 135765

QY 601 aatgtgggtctcctggatcagaggctggccctgcagtgggtgcaggagaacgtggcagcc 660
|
Db 135764 AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 135705

QY 661 ttcgggggtgaccccgacatcagtgacgctgtttggggagagcgcgggagccgctcggtg 720
|
Db 135704 TTCGGGGGTGACCCCGACATCAGTGACCGCTGTTTGGGGAGAGCGCGGGAGAGCCGCTCGGTG 135645

QY 721 ggcattgcacctgctgtccccgccccagccggggccctgttccacagggcgctgctgcagagc 780
|
Db 135644 GGCATGCACCTGCTGTCCCCGCCAGCCCGGGCCCTGTTCACAGGGCCGTGCTGCAGAGC 135585

QY 781 ggtgcccccaatggaccctggccacggtgggcatgggagagggccccgcagggccacg 840
|
Db 135584 GGTGCCCCCAATGGACCCCTGGGCCACGTTGGGCATGGGAGAGGCCCGTCGAGGGCCACG 135525

QY 841 cagctggcccaccttgtgggctgtcctcaggcggcactggtgggaatgacacagagctg 900
|
Db 135524 CAGCTGGCCCACTTGTGGGCTGTCTCCAGCGGGCACTGGTGGGAATGACACAGAGCTG 135465
```


QY	901	gtagcctgccttcggacacgaccagcgaggctccttggtgaaccacgaatggcacgtgctg	960
Db	135464	GTAGCCTGCCTTTCGGACACGACGACGCGAGGTCCTGGTGAACCACGAATGGCACGTGCTG	135405
QY	961	cctcaagaagcgtcttcctccggttctccttccttcgtgcctgtgtagatggagacttcctcagt	1020
Db	135404	CCTCAAGAAAGCGTCTTCCGTTCTCTCTTCGTGCCTGTGTAGATGGAGACTTCCTCAGT	135345
QY	1021	gacacccacagagggccctcatcaacgcgggagacattccacggcctg-----	1065
Db	135344	GACACCCACAGAGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTAAC TAGTGGC	135285
QY	1066	-----	1065
Db	135284	TAGTGGCGTGAAGCTGGCTCCTCTGTGGTCCCAACGGTCCCTCCCTCGCAGGGACC	135225
QY	1066	-----	1065
Db	135224	CAGGCATGAGGGCTTCTCCAGGCCCATTCCCAGAAGTCCCAAGATCCTCCCTGAGGGCT	135165
QY	1066	-----	1065
Db	135164	CAGATCCCAGGGTGGTCAGCAGGGCAGACAGGAAAGCCACCATGGGTCTATTTCTCTTT	135105
QY	1066	-----	1065
Db	135104	CTCTGCATCCCTCCCCCTGATCTCGTCTCTCTCTGTCCATGTTCCGGGTCTGTAACTGT	135045
QY	1066	-----	1065
Db	135044	TCATCTCTCTGGCTCTTTGTCTGTCCATCTGTTTCTGTCTACTTGTCTGTGTCCTGT	134985
QY	1066	-----caggtgctgggtggtgtgtgtgaaggatga	1094
Db	134984	CGCTCCATCCCACCCCTCTCCCTCACCCCGAGGTGCTGGTGGTGGTGAAGGATGA	134925
QY	1095	gggctcgattttctgttttaagggggcccccaggcttcagaaagacaacgagtcctcat	1154
Db	134924	GGGCTCGTATTTTCTGGTTTACGGGGCCCCCAGGCTTCAGCAAGACAAACGAGTCTTCAT	134865
QY	1155	cagccgggcccagttcctcgccgggggtgcgggtcggggtcccccaggttaagtgcacctggc	1214
Db	134864	CAGCCGGGCCGAGTTCCTTGCCGGGGTGCGGGTTCGGGTTCCTCCAGGTAAGTGACCTGGC	134805
QY	1215	agccaggctgtggtcctgcattacacagactggctgcatacccgaggaccggcacgcct	1274
Db	134804	AGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATGCCGAGGACCCGGCACGCCT	134745
QY	1275	gagggagggcctgagcgaatgtggtggcgaccacaatgtcgtgtgccccgtggccccagct	1334
Db	134744	GAGGAGGGCCCTGAGCGATGTGTGGGGCGACCAACAATGTCTGTGCCCCGTGGCCCCAGCT	134685
QY	1335	ggctgggcgactggctgccccagggtgccccgggtctacgcctacgtctttgaacacccgtgc	1394
Db	134684	GGCTGGGGCGACTGGCTGCCCCAGGGTGCCCCGGGTCTACGGCTACGTCTTTGAACACCCGTGC	134625
QY	1395	tccacgcctcctcgccccctgtggatgggggtgccccacggctacgagatcgagttcat	1454
Db	134624	TTCCACGCTCTCTTGCCCCCTGTGGATGGGGTGGCGGTACGGCTACGGCTACGAGATCGAGTTTCA	134565
QY	1455	ctttgggataccccctggaccctctcgaactacacggcagaggagaaaaatcttcgccca	1514
Db	134564	CTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAGAAATCTTCGCCCA	134505
QY	1515	gcgactgatcgatactgggccaactttgcccccacaggg	1554
Db	134504	CGGACTGATCGGATACTGGGCCCAACTTTGCCCGCACAGGG	134465
RESULT	14		
AC084057			
LOCUS	AC084057	194681 bp	DNA linear HTG 12-OCT-2000

DEFINITION	Homo sapiens chromosome 7 clone CTC-786K12, WORKING DRAFT SEQUENCE, 36 unordered pieces.
ACCESSION	AC084057
VERSION	AC084057.1 GI:10799406
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 194681)
TITLE	DOE Joint Genome Institute.
JOURNAL	Sequencing of Human Chromosome 7 Unpublished
REFERENCE	2 (bases 1 to 194681)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (12-OCT-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	-----Genome Center Center: Joint Genome Institute Center Code: JGI Web site: http://www.jgi.doe.gov ----- Project Information Center Project Name: 2351286 Center clone name: CTC-786K12 ----- Summary Statistics Consensus quality: 162650 bases at least Q40 Consensus quality: 173899 bases at least Q30 Consensus quality: 176608 bases at least Q20 Estimated insert size: 200000; pulse field gel estimation Estimated insert size: 191181; sum-of-contigs estimation Quality coverage: 7.95 in Q20 bases; pulse field gel estimation Quality coverage: 8.32 in Q20 bases; sum-of-contigs estimation. * NOTE: This is a 'working draft' sequence. It currently * consists of 36 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * 1 1324: contig of 1324 bp in length * 1325 1424: gap of unknown length * 1425 2586: contig of 1162 bp in length * 2587 2686: gap of unknown length * 2687 4056: contig of 1370 bp in length * 4057 4156: gap of unknown length * 4157 5197: contig of 1041 bp in length * 5198 5297: gap of unknown length * 5298 6413: contig of 1116 bp in length * 6414 6513: gap of unknown length * 6514 7675: contig of 1162 bp in length * 7676 7775: gap of unknown length * 7776 8843: contig of 1068 bp in length * 8844 8943: gap of unknown length * 8944 9998: contig of 1055 bp in length * 9999 10098: gap of unknown length * 10099 11598: contig of 1500 bp in length * 11599 11698: gap of unknown length * 11699 12956: contig of 1258 bp in length * 12957 13056: gap of unknown length * 13057 14168: contig of 1112 bp in length * 14169 14268: gap of unknown length * 14269 15287: contig of 1019 bp in length * 15288 15387: gap of unknown length * 15388 16936: contig of 1549 bp in length * 16937 17036: gap of unknown length * 17037 18086: contig of 1050 bp in length * 18087 18186: gap of unknown length * 18187 19427: contig of 1241 bp in length * 19428 19527: gap of unknown length

*	19528	20570:	contig of 1043 bp in length	
	20571	20670:	gap of unknown length	
*	20671	22193:	contig of 1523 bp in length	
*	22194	22293:	gap of unknown length	
*	22294	24473:	contig of 2180 bp in length	
*	24474	24573:	gap of unknown length	
*	24574	25995:	contig of 1422 bp in length	
*	25996	26095:	gap of unknown length	
*	26096	27843:	contig of 1748 bp in length	
*	27844	27943:	gap of unknown length	
*	27944	30120:	contig of 2177 bp in length	
*	30121	30220:	gap of unknown length	
*	30221	31961:	contig of 1741 bp in length	
*	31962	32061:	gap of unknown length	
*	32062	35719:	contig of 3658 bp in length	
*	35720	35819:	gap of unknown length	
*	35820	38071:	contig of 2252 bp in length	
*	38072	38171:	gap of unknown length	
*	38172	41171:	contig of 3000 bp in length	
*	41172	41271:	gap of unknown length	
*	41272	47552:	contig of 6281 bp in length	
*	47553	47652:	gap of unknown length	
*	47653	52017:	contig of 4365 bp in length	
*	52018	52117:	gap of unknown length	
*	52118	55609:	contig of 3492 bp in length	
*	55610	55709:	gap of unknown length	
*	55710	60279:	contig of 4570 bp in length	
*	60280	60379:	gap of unknown length	
*	60380	66246:	contig of 5867 bp in length	
*	66247	66346:	gap of unknown length	
*	66347	78184:	contig of 11838 bp in length	
*	78185	78284:	gap of unknown length	
*	78285	96184:	contig of 17900 bp in length	
*	96185	96284:	gap of unknown length	
*	96285	115729:	contig of 19445 bp in length	
*	115730	115829:	gap of unknown length	
*	115830	139049:	contig of 23220 bp in length	
*	139050	139149:	gap of unknown length	
*	139150	161902:	contig of 22753 bp in length	
*	161903	162002:	gap of unknown length	
*	162003	194681:	contig of 32679 bp in length.	
FEATURES				
source				
1..194681				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/chromosome="7"				
/clone="CTC-786K12"				
/clone_lib="CalTech human BAC library C"				
BASE COUNT 45257 a 49777 c 50799 g 45331 t 3517 others				
ORIGIN				
Query Match 69.4%; Score 1198; DB 2; Length 194681;				
Best Local Similarity 81.8%; Pred. No. 3.6e-202;				
Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;				
QY	1	atgagccccgcagtgctgctgcacagccttccctggcttcccaactccttctctc	60	
Db	106067	ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCCTTCCCTGGCTTCCCACTCTCTCCTC	106126	
QY	61	ctcctctggtcctgggtggaggagtggggctgagggccgggagatgcagagctgctg	120	
Db	106127	CTCCTCTGGCTCTCTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG	106186	
QY	121	gtgacggtgcgtggggccggctgcggggcattcgcctgaagacccccggggccctgtc	180	
Db	106187	GTGACGGTGCgtggggccggctgcggggcattcgcctgaagacccccggggccctgtc	106246	
QY	181	tctgctttctgggcatccccctttgcggagccaccatgggaccccgcttcttgcca	240	
Db	106247	TCTGCTTCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGCTTCTGCCA	106306	
QY	241	ccggagcccaagcagccttggtcaggggtggtagacgcctacaaccttccagagtgtctgc	300	

Db	106307	CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC	106366	
QY	301	taccaatatgtggacaccctataccccaggttttgaggccaccgagatgtggaacccaac	360	
Db	106367	TACCAATATGTGGACACCCTATACCCAGSTTTTGAGGGCACCGAGATGTGGAACCCCAAC	106426	
QY	361	cgtgagctgagcaggactgcctgtacctcaacgtgtggacaccatacccccgccctaca	420	
Db	106427	CGTGAGCTGACGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCGGCCCTACA	106486	
QY	421	tccccacccctgtcctcgtctcgtgatctatgggggtggttcttacagtgggcctctcc	480	
Db	106487	TCCCCACCCCTGTCTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCTCC	106546	
QY	481	ttggacgtgtacgatggccgcttcttggtaacaggccgagaggaactgtctggtgtccatg	540	
Db	106547	TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG	106606	
QY	541	aactaccgggtgggagcccttggcttctcctggccctgccgggagccgagagagccccgggc	600	
Db	106607	AACTACCGGTGGGAGCCTTTGGCTTCTCTGGCCTTCCGGGAGCCGAGAGGCCCGGGGC	106666	
QY	601	aatgtgggtcctcctggatcagaggctggccctgcagtggtgcaggagaacgtggcagcc	660	
Db	106667	AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCC	106726	
QY	661	ttcgggggtgaccccgacatcagtgacgctgttttggggagagcgcgggagccgctcggtg	720	
Db	106727	TTCGGGGGTGACCCGACATCAGTGACCGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG	106786	
QY	721	ggcatgcacctgctgtcccccggccacggtgggcatgggagggccctgcagagcagc	780	
Db	106787	GGCATGCACCTGTGTCTCCGCCAGCCGGGCTGTTCACAGGGCGCTGTCTGCAGAGC	106846	
QY	781	ggtgcccccaatggaccctggccacggtgggcatgggagggccctgcagggccacg	840	
Db	106847	GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCTCGCAGGGCCACG	106906	
QY	841	cagctggcccaacctgtggctgtcctcagggcgccactggtgggaatgacacagagctg	900	
Db	106907	CAGCTGGCCCACTTGTGGCTGTCTCTCCAGCGGCACCTGGTGGGAATGACACAGAGCTG	106966	
QY	901	gtagcctgccttcggacacgaccagcgcaggtcctggtgaaccacgaatggcacgtgctg	960	
Db	106967	GTAGCCTGCCTTCGGACACGACCAGCGCAGGTCTCTGGTGAACCCACGAATGGCACGTGCTG	107026	
QY	961	cctcaagaaagcgtcttcccggttctcctcgtgcctgtggtagatggagacttccctcagt	1020	
Db	107027	CCTCAAGAAAGCGTCTTCCGGTTCTCTCTCGTGCCTGTGGTAGATGGAGACTTCCTCAGT	107086	
QY	1021	gacacccccagagggccctcatcaacgcgggagacttccacggcctg	1065	
Db	107087	GACACCCCAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTAACTAGTGGC	107146	
QY	1066	-----	1065	
Db	107147	TAGCTGGCGTGAAGCTGGCTCCTCTGGGTCCCAACGGTCCCTCCCTGCCCTGCAGGGACC	107206	
QY	1066	-----	1065	
Db	107207	CAGGCATGAGGGCTTCTCCAGGCCCATTTCCAGAGAAGTCCCAGAAGTCTCTCCCTGAGGGCT	107266	
QY	1066	-----	1065	
Db	107267	CAGATCCCAGGGTGGTTCAGAGGGCAGACAGGAAGCCACCATGGGTCTATTTTCTCTTT	107326	
QY	1066	-----	1065	
Db	107327	CTCTGCATCCCTCCCTGATCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	107386	
QY	1066	-----	1065	

Db 107387 TCATCTCTCTGGCTCTTTGTGTCGTGCCATCTGTGTTTCTGTCTACTTGTCTGTCTGTGCCTGT 107446

QY 1066 -----caggtgctggtgggtggtgaaggatga 1094
|||||

Db 107447 CGTCCATCCACACCCCTCTCCCTACCCCCAGGTGCTGGTGGTGGTGAAGGATGA 107506

QY 1095 gggctcgattttctggtttacggggcccgaggttcagcaaaagacaacgagtcctctcat 1154
|||||

Db 107507 GGGCTCGATTTTCTGTTTACGGGGCCCCAGGCTTCAGCAAAAGACAACGAGTCTCTCAT 107566

QY 1155 cagccgggscgagttcctgcccggggtcggggtcggggttccccaggttaagtgcctggc 1214
|||||

Db 107567 CAGCCGGSCGAGTTCTTGCCCGGGTTCGGGGTTCGGGGTTCCCCAGGTAAGTGACCTGGC 107626

QY 1215 agccgagctggtggtcctgcattacacagactggctgcatacccgaggaccccgcacgcct 1274
|||||

Db 107627 AGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCGAGGACCCGGCACGCCT 107686

QY 1275 gagggagggcctgagcgatgtggtggcgagaccacaatgtcgtgcccccggtgcccagct 1334
|||||

Db 107687 GAGGGAGGCCCTGAGCGATGTGGTGGGCGACCAACAATGTCTGTGCCCCGTGGCCCCAGCT 107746

QY 1335 ggctgggcactggtgcccaggggtgcggggtctacgcctacgtctttgaacacccgtgc 1394
|||||

Db 107747 GGCTGGGCGACTGGCTGCCCAGGCTGCCGGGTCTACGCCCTACGCTCTTTGAACACCGTGC 107806

QY 1395 ttccacgctctcctgccccctgtggatgggggtgccccacggctacgagatcgagttcat 1454
|||||

Db 107807 TTCCACGCTCTCTGTGGCCCCCTGTGGATGGGGGTGCCCCACGGCTACGAGATCGAGTTTCAT 107866

QY 1455 ctttgggatcccccctggaccccctctcgaaactacacggcgagagagaaaaattcttcgccca 1514
|||||

Db 107867 CTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAGAGAAATCTTCGCCCA 107926

QY 1515 gcgactgatcgatactgggccaactttgcccgcacaggg 1554
|||||

Db 107927 GCGACTGATCGGATACTGGGCCAACTTTGCCCCGCACAGGG 107966

RESULT 15
HSAF002993/c

LOCUS HSAF002993 34921 bp DNA linear PRI 12-JUN-2000

DEFINITION Homo sapiens cosmid from 7q22, complete sequence.

ACCESSION AF002993

VERSION AF002993.1 GI:2735699

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 34921)

AUTHORS Shapira,M., Tur-Kaspa,I., Bosgraaf,L., Livni,N., Grant,A.D., Grisar,D., Korner,M., Ebstein,R.P. and Soreq,H.

TITLE A transcription-activating polymorphism in the ACHE promoter associated with acute sensitivity to anti-acetylcholinesterases

JOURNAL Hum. Mol. Genet. 9 (9), 1273-1281 (2000)

MEDLINE 20275616

PUBMED 10814709

REFERENCE 2 (bases 1 to 34921)

AUTHORS Drescher,B., Rosenthal,A., Hinzmann,B. and Korner,M.

TITLE Genomic sequencing in 7q22 revealed a novel arsenite resistance gene

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 34921)

AUTHORS Drescher,B., Rosenthal,A., Hinzmann,B. and Korner,M.

TITLE Direct Submission

JOURNAL Submitted (07-MAY-1997) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

FEATURES

source

1. .34921
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"

/map="7q22"
804. .973
/note="MZEf, score = 99.9%"
/evidence=not_experimental
809. .973
/note="Xpound exon prediction, score = 96% (0%)"
/evidence=not_experimental
841. .1011
/note="BLASTN (EMBL), 103. .273 of emb|G30916|HS1140102 human chromosome 7 STS SWSS3457; single read."
/note="Region: Data base match"
/evidence=not_experimental
1102. .1286
/note="MZEf, score = 99.9%"
/evidence=not_experimental
1105. .1286
/note="Xpound exon prediction, score = 96% (0%)"
/evidence=not_experimental
1474. .1572
/note="Xpound exon prediction, score = 92% (0%)"
/evidence=not_experimental
1474. .1572
/note="MZEf, score = 100%"
/evidence=not_experimental
1657. .1774
/note="MZEf, score = 99.8%"
/evidence=not_experimental
1657. .1774
/note="Xpound exon prediction, score = 97% (0%)"
/evidence=not_experimental
1947. .2104
/note="Xpound exon prediction, score = 98% (0%)"
/evidence=not_experimental
1947. .2104
/note="MZEf, score = 100%"
/evidence=not_experimental
2005. .2119
/note="CpG island score = 0.87, GC = 67.80%, CpGs = 13"
/note="Region: CpG island"
/evidence=not_experimental
2105. .2434
/note="Alignment: EST, | Intron"
/evidence=not_experimental
2497. .2596
/note="Alignment: EST, | Intron"
/evidence=not_experimental
2597. .2675
/note="Xpound exon prediction, score = 98% (0%)"
/evidence=not_experimental
2597. .2675
/note="MZEf, score = 98.9%"
/evidence=not_experimental
2676. .2980
/note="Alignment: EST, | Intron"
/evidence=not_experimental
3170. .3528
/note="Alignment: EST, | Intron"
/evidence=not_experimental
3529. .3585
/note="MZEf, score = 100%"
/evidence=not_experimental
3539. .3764
/note="CpG island score = 0.82, GC = 62.80%, CpGs = 18"
/note="Region: CpG island"
/evidence=not_experimental
3548. .3702
/note="Xpound exon prediction, score = 98% (0%)"
/evidence=not_experimental
3637. .3702
/note="Alignment: CDS, |CG415001|87.9|46|111"
/evidence=not_experimental
3703. .3780
/note="Alignment: CDS, | Intron"

exon /evidence-not_experimental
3781. .3934
/note="MZF, score = 97.3%"
/evidence-not_experimental
exon 3781. .3934
/note="Xpound exon prediction, score = 99% (0%)"
/evidence-not_experimental
exon 3781. .3927
/note="Alignment: CDS, |CG415001|85.7|112|258"
/evidence-not_experimental
misc_feature 3812. .3960
/note="BLASTN (EMBL), 110. .258 of emb|U41500|CG41500
Cricetulus griseus arsenite-resistance protein (ars2)
mRNA,"
/note="Region: Data base match"
/evidence-not_experimental
exon 4054. .4281
/note="MZF, score = 84.8%"
/evidence-not_experimental
exon 4054. .4247
/note="Xpound exon prediction, score = 98% (0%)"
/evidence-not_experimental
exon 4191. .4220
/note="BLASTX (GENPEPT), 58. .87 of gb|U41500|CGU41500_1
Cricetulus griseus arsenite-resistance protein (ars2)"
/evidence-not_experimental
intron 4248. .4436
/note="Alignment: CDS, | Intron"
/evidence-not_experimental
exon 4437. .4593
/note="MZF, score = 97%"
/evidence-not_experimental
exon 4437. .4593
/note="Xpound exon prediction, score = 89% (0%)"
/evidence-not_experimental
exon 4437. .4593
/note="Alignment: CDS, |CG415001|86.6|457|613"
/evidence-not_experimental
misc_feature 4466. .4626
/note="BLASTN (EMBL), 453. .613 of emb|U41500|CG41500
Cricetulus griseus arsenite-resistance protein (ars2)
mRNA,"
/note="Region: Data base match"
/evidence-not_experimental
exon 4467. .4519
/note="BLASTX (GENPEPT), 87. .139 of gb|U41500|CGU41500_1
Cricetulus griseus arsenite-resistance protein (ars2)"
/evidence-not_experimental
intron 4594. .4788
/note="Alignment: EST, | Intron"
/evidence-not_experimental
exon 4809. .4840
/note="BLASTX (GENPEPT), 140. .171 of gb|U41500|CGU41500_1
Cricetulus griseus arsenite-resistance protein (ars2)"
/evidence-not_experimental
intron 4879. .4991
/note="Alignment: CDS, | Intron"
/evidence-not_experimental
exon 4992. .5118
/note="Alignment: CDS,
|CG415001|88.3|717|844|CG415002|88.6|1144"
/evidence-not_experimental
misc_feature 5045. .5152
/note="BLASTN (EMBL), 738. .845 of emb|U41500|CG41500
Cricetulus griseus arsenite-resistance protein (ars2)
mRNA,"
/note="Region: Data base match"
/evidence-not_experimental
intron 5119. .5208
/note="Alignment: CDS, | Intron"
/evidence-not_experimental
exon 5209. .5264
/note="Alignment: CDS, |CG415001|89.3|845|900"

misc_feature /evidence-not_experimental
complement(5238. .5431)
/note="BLASTN (EMBL), 1. .194 of emb|G30851|HS1140038 human
chromosome 7 STS sWSS3840; single read."
/note="Region: Data base match"
/evidence-not_experimental
misc_feature 5757. .14176
/note="GC score = 489.57 (8420bp)"
/note="Region: GC content"
/evidence-not_experimental
misc_feature 5844. .7208
/note="CpG island score = 0.79, GC = 69.50%, CpGs = 126"
/note="Region: CpG island"
/evidence-not_experimental
misc_feature 6019. .6471
/note="CpG core score = 8.90"
/note="Region: CpG island"
/evidence-not_experimental
repeat_region 6454. .6508
/note="homology = 89.10%, score = 38, counts = 11"
/rpt_type=tandem
/rpt_unit=cctcc
Query Match 65.3%; Score 1126; DB 9; Length 34921;
Best Local Similarity 81.2%; Pred. NO. 2.5e-189;
Matches 1505; Conservative 0; Mismatches 0; Indels 349; Gaps 3;
QY 1 atgagggccccgcagtgtctgtgcacacgccttcctggcttccccactccttcctc 60
|||||
Db 10809 ATGAGGCCCCCGCAGTGTCTGTGTGCACACGCCCTTCCTGGCTTCCCCACTCTTCTCCTC 10750
QY 61 ctcctctggctcctgggtggaggagtggggctgagggccgggagagatgcagagctgtg 120
|||||
Db 10749 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 10690
QY 121 gtgacggtgcgtggggccggctgcggggcattccctgaagacccccggggccctgtc 180
|||||
Db 10689 GTGACGGTGCgtggggccggctgcggggcattccctgaagacccccggggccctgtc 10630
QY 181 tctgcttccctgggcattcccttgcggagccaccatgggaccccgctgttcttccca 240
|||||
Db 10629 TCTGCTTTCCTGGGCATCCCTTTTGGGAGCCACCCATGGGACCCCGTCGCTTCTTGCCA 10570
QY 241 ccggagcccaagcagccttggtcaggggtggtagcgcctacaaccttccagagtgtctgc 300
|||||
Db 10569 CCGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGTACAACCTTCCAGAGTGTCTGC 10510
QY 301 taccaatatgtggacaccctataccagggtttttaggggcaccgagatgtggaaccccaac 360
|||||
Db 10509 TACCAATATGTGGACACCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 10450
QY 361 cgtgagctgagcaggactgcctgtacctcaacgtgttgacaccatacccccgccctaca 420
|||||
Db 10449 CGTGAGCTGAGCGAGACTGCCTGTACCTAACGTTGTGGACACCATAACCCCGCCCTACA 10390
QY 421 tccccacccctgtcctcgtctgtgatctatgggggtggcttctacagtggggcctcctcc 480
|||||
Db 10389 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCCTCC 10330
QY 481 ttggacgtgtacgatggccgcttcttggtacaggccgagagagactgtgctggtgtccatg 540
|||||
Db 10329 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 10270
QY 541 aactaccgggtgggaccccttggcttcctggccctgcccgggagccgagagccccgggc 600
|||||
Db 10269 AACTACCGGGTGGGAGCCTTTGGCTTCCCTGGCCCTGCCGGGAGCCCGCCGGGC 10210
QY 601 aatgtgggtctcctggtacagaggtggccctgagctgggtgaggaagacgtggcagcc 660
|||||
Db 10209 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGCAGCC 10150
QY 661 ttcgggggtgacccacatcagtgacgctgtttgggagagcgcgggagccgctcggtg 720
|||||

Db 10149 TTCCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 10090

Qy 721 ggcattgacacctgctgtcccccgcagccgggacctgttccacagggccgtgctgcagagc 780
|||||

Db 10089 GGCATGACACTGCTGTCCCCGCCAGCCGGGCTGTTCACAGGGCCGTGCTGCAGAG- 10029

Qy 781 ggtgcccccaatggaccctgggccaaggtgggcatgggagaggcccgccgagggccacg 840
|||||

Db 10030 -----CGCAGGGCCACG 10019

Qy 841 cagctggccccaccttgtgggctgtcctccaggcggcactggtgggaatgacacagagctg 900
|||||

Db 10018 CAGCTGCCCCACCTTGTGGCTGTCTCCAGGGCCACTGGTGGGAATGACACAGAGCTG 9959

Qy 901 gtagcctgccttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
|||||

Db 9958 GTAGCCTGCCTTCGGACACGACCGACGAGGTCTTGGTGAACCAAGCATGGCAGCTGCTG 9899

Qy 961 cctcaagaaagcgtcttccggttctccttgcctgctggtagatggagacttcctcagt 1020
|||||

Db 9898 CCTCAAGAAAGCGTCTTCCGGTTCCTCTCGTGCCTGTGGTAGATGGAGACTTCCTCAGT 9839

Qy 1021 gacacccagagggcctcatcaacgggggagacttccacggcctg----- 1065
|||||

Db 9838 GACACCCAGAGGCCCTCATCAACGGGGGAGACTTCCACGGCCTGCAGGTAAGTAGTGGC 9779

Qy 1066 ----- 1065

Db 9778 TAGCTGGCGTGAAGCTGGCTCCTCTGGGTCCCAACGGTCCCTCCCTCCCTGCAGGGACC 9719

Qy 1066 ----- 1065

Db 9718 CAGGCATGAGGGCTTCTCCAGGCCCATGTTCAGCAGGGCAGACAGGAAAGCCACCATGGGT 9659

Qy 1066 ----- 1065

Db 9658 CTATTTTCTCTTCTCTGCATCCCTCCCTGATCTCGTCTCTCTCTGTCCATGGTTCCG 9599

Qy 1066 ----- 1065

Db 9598 GGCTGTAACTGTTTCATCTCTCTGGCTCTTGTCTGTCCATCTGTTTGTCTACTTGTG 9539

Qy 1066 -----caggtgctggtgggtg 1081
|||||

Db 9538 TGTCTGTGCTGTGCTGCCATCCCAACCCCCCTCTCCCTCACCCCCAGGTGCTGGTGGGTG 9479

Qy 1082 tggtagaggatgagggctcgtattttctggttttacgggggccccaggcttcagcaaaagaca 1141
|||||

Db 9478 TGGTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAGACA 9419

Qy 1142 acgagtctctcatcagccgggcccagttcctgcccgggtgcgggtgcgggttccccagg 1201
|||||

Db 9418 ACGAGTCTCTCATCAGCCGGGCGGAGTCTCTGGCCGGGTGCGGGTTCGCCAGG 9359

Qy 1202 taagtacctggcagccgaggtgtggtcctgcattacacagactggtgctgcatcccgagg 1261
|||||

Db 9358 TAAGTGACCTGGCAGCCGAGGCTGTGTCTGTCATTACACAGACTGGCTGCATCCCCGAGG 9299

Qy 1262 acccggcacgcctgaggagggccctgagcgtgtggtggcgaccacaatgtcgtgtgcc 1321
|||||

Db 9298 ACCCGGCACGCCTGAGGGAGGCCCTGAGCGATGTGTGGCGCACCAATGTCTGTGTGCC 9239

Qy 1322 ccgtggcccagctggctgggagactggctgcccc-aggggtccccgggtctacgcctacgtc 1380
|||||

Db 9238 CCGTGGCCCCAGCTGGCTGGGCGACTGGCTGCCCCNAGGGTGCCCGGGTCTACGCCCTACGTC 9179

Qy 1381 ttgtaacacogtgccttccacgctctcctggccccctgtggatgggggtgccccacggctac 1440
|||||

Db 9178 TTTGAACACCGTGTCTCCACGCTCTCTGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 9119

Qy 1441 gagatcgagttcatctttgggataccccctggacccctctcgaactacacggcagaggag 1500
|||||

Db 9118 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTTCGAAACTACACGGCAGAGGAG 9059

Qy 1501 aaaatcttcgccccagcgactgatgcgatactgggccaactttgccccgcacaggg 1554
|||||

Db 9058 AAAATCTTGGCCCCAGCGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGG 9005

Search completed: August 31, 2002, 18:18:14
Job time: 16547 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:25:58 ; Search time 824.08 Seconds
(without alignments)
3593.920 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atgaggccccgcagtgctt.....aattgctcagcgccaccta 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_032802:*	
	1:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT:*			
	2:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT:*			
	3:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT:*			
	4:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT:*			
	5:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT:*			
	6:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT:*			
	7:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT:*			
	8:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT:*			
	9:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT:*			
	10:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT:*			
	11:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT:*			
	12:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT:*			
	13:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT:*			
	14:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT:*			
	15:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT:*			
	16:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT:*			
	17:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT:*			
	18:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT:*			
	19:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT:*			
	20:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT:*			
	21:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT:*			
	22:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT:*			
	23:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT:*			
	24:	/SIDS1/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1725	100.0	1725	24	AAS17549
2	1722	99.8	1845	24	AAS17493
3	1722	99.8	2256	16	AAQ99002
4	1718.6	99.6	5767	24	AAS17547
5	1718.6	99.6	14446	24	AAS17548
6	1684.8	97.7	2253	11	AAQ05998
7	1313.6	76.2	1800	11	AAQ05999
8	1282	74.3	2089	24	ABI99790
9	1198	69.4	12113	24	AAS17492
					Synthetic human ac
					Human cDNA encodin
					Human acetylcholin
					Plasmid pTM034. S
					Plasmid pTM036. S
					Sequence encoding
					Mouse ischaemic co
					Human acetylcholin

10	741.8	43.0	1698	20	AAX86613	CDNA encoding an a
11	741.8	43.0	2478	20	AAX86614	CDNA encoding an a
12	380	22.0	2416	21	AAZ49470	Human wild type Bu
13	380	22.0	2416	21	AAZ49471	Human Butyrylcholi
14	378.4	21.9	2400	14	AAQ42496	Full-length human
15	322.8	18.7	2445	7	AAN60111	Sequence encoding
16	322.8	18.7	2445	14	AAQ48393	Full-length foetal
17	200.4	11.6	4377	23	ABL19443	Drosophila melanog
18	185.6	10.8	764	7	AAN60110	Sequence encoding
19	185.6	10.8	764	14	AAQ42495	Foetal brain clone
20	142.6	8.3	1614	16	AAQ90569	Recombinant bile s
21	142.6	8.3	1614	17	AAT47862	Human wild-type bi
22	142.6	8.3	2166	16	AAQ90579	Encodes bile salt-
23	142.6	8.3	2166	17	AAT47861	CDNA of human bile
24	142.6	8.3	3018	17	AAT34787	Human milk bile sa
25	141	8.2	2428	13	AAQ20252	Bile-salt stimulat
26	141	8.2	3018	12	AAQ14186	Human milk bile-sa
27	139.4	8.1	2184	15	AAQ71874	Sequence encoding
28	139.4	8.1	2344	22	AAF31110	Cholesterol estera
29	139.4	8.1	2428	17	AAT35250	Human bile salt-st
30	139.4	8.1	2428	20	AAZ20300	Human bile salt-st
31	139.4	8.1	2428	24	ABA04403	Human milk bile sa
32	139.4	8.1	2428	24	AAS18538	CDNA encoding Bile
33	139.4	8.1	2487	21	AAA63883	CDNA encoding a hu
34	137.4	8.0	2428	15	AAQ83221	Bile salt-stimulat
35	130.4	7.6	1908	12	AAQ11977	Bovine pancreatic
36	122.6	7.1	2520	22	AAH57518	Human pancreas cel
37	120.2	7.0	1680	22	AAF25258	Nucleotide sequenc
38	120.2	7.0	2202	23	AAZ74460	DNA encoding novel
39	120.2	7.0	2484	21	AAF21884	Human breast and o
40	118.2	6.9	2076	22	AAD08430	Human secreted pro
41	118.2	6.9	27314	23	ABL19442	Drosophila melanog
42	117.4	6.8	2101	21	AAD00680	Human Hydrolase pr
43	117.4	6.8	3776	22	AAD08405	Human secreted pro
44	117.4	6.8	3824	21	AAA96341	CDNA encoding a no
45	117.4	6.8	3824	22	AAS21514	Human cDNA sequenc

ALIGNMENTS

RESULT 1	
AAS17549	
ID	AAS17549 standard; DNA; 1725 BP.
XX	
AC	AAS17549;
XX	
DT	25-FEB-2002 (first entry)
XX	
DE	Synthetic human acetylcholinesterase gene.
XX	
KW	Human; acetylcholinesterase; AChE; antidote; pesticide;
KW	transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW	muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW	organophosphate(OP)-modified AChE; pyridostigmine bromide; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	WO200171014-A2.
XX	
PD	27-SEP-2001.
XX	
PF	16-MAR-2001; 2001WO-US08468.
XX	
PR	17-MAR-2000; 2000US-190440P.
XX	
PA	(MORT/) MOR T.
PA	(SORE/) SOREQ H.
PA	(ARNT/) ARNTZEN C.
PA	(MASO/) MASON H.
PA	(BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX	

|||||
Db 1561 aatgagccccgagaccccaaggccccacaatgcccccgtagacggcggggctcagcag 1620
Qy 1621 taegttagtctggacctgcggccgctggaggtgcggcggggctgcgcgccagcctgc 1680
|||||
Db 1621 taegttagtctggacctgcggccgctggaggtgcggcggggctgcgcgccagcctgc 1680
Qy 1681 gccttctggaaccgcttctcccaaatgtctcagcgccacctga 1725
|||||
Db 1681 gccttctggaaccgcttctcccaaatgtctcagcgccacctga 1725
RESULT 2
AAS17493
ID AAS17493 standard; cDNA; 1845 BP.
XX
AC AAS17493:
XX
DT 26-FEB-2002 (first entry)
XX
DE Human cDNA encoding acetylcholinesterase, ACHE.
XX
KW Human; ss; ACHE; acetylcholinesterase; polymorphic variant; haplotyping;
KW genotyping; neurological disease; Parkinson's disease;
KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..1845
FT /*tag= a
FT /product= "ACHE"
FT replace (36,T)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT replace (101,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT replace (183,C)
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT replace (838,G)
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
FT replace (873,T)
FT /*tag= f
FT /standard_name= "Single nucleotide polymorphism"
FT replace (1032,A)
FT /*tag= g
FT /standard_name= "Single nucleotide polymorphism"
FT replace (1057,A)
FT /*tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT replace (1431,T)
FT /*tag= i
FT /standard_name= "Single nucleotide polymorphism"
XX
PN WO200179219-A2.
XX
PD 25-OCT-2001.
XX
PF 11-APR-2001; 2001WO-US11853.
XX
PR 14-APR-2000; 2000US-197173P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
PA (KAZE/) KAZEMI A.
XX
PI Bentivegna SC, Chew A, Choi JY, Koshy B;
XX
DR WPI; 2002-055248/07.
DR P-PSDB; AAU11231.
XX

PT New polymorphic variants comprising acetylcholinesterase (ACHE)
isogene, useful in expressing ACHE protein for use in screening for
PT candidate drugs to treat diseases related to ACHE activity, e.g.
PT neurological diseases or cancer -
XX
PS Claim 26; Fig 2; 79pp; English.
XX
CC The invention relates to a polynucleotide comprising a polymorphic
variant of an acetylcholinesterase (ACHE) gene or fragment, protein or
CC complement, the variant comprising an ACHE isogene defined by a haplotype
CC selected from haplotypes 1-20 listed in the specification. Also included
CC are methods for haplotyping and genotyping the ACHE gene of an
CC individual, a method for predicting a haplotype pair for the ACHE gene of
CC an individual, a method for identifying an association between a trait
CC and at least one haplotype or haplotype pair of ACHE gene, recombinant
CC nonhuman organisms transformed or transfected with the polynucleotide
CC where the organism expresses ACHE protein encoded by the first
CC nucleotide sequence or encoded by the polymorphic variant sequence,
CC an isolated antibody specific for and immunoreactive with ACHE,
CC a method of screening for drugs targeting the polypeptide contacting ACHE
CC polymorphic variant with a candidate agent and assaying for binding
CC activity, a computer system for storing and analysing polymorphism data
CC for ACHE gene and a genome anthology for ACHE gene which comprises ACHE
CC isogenes defined by haplotypes 1-20 given in the specification.
CC The Polymorphisms are useful for studying the biological function of
CC ACHE as well as in identifying drugs targeting this protein for the
CC treatment of disorder related to its abnormal expression or function.
CC The polymorphic variants may also be used in screening for compounds
CC targeting ACHE to treat a specific condition or disease predicted to be
CC associated with ACHE activity e.g. neurological diseases (e.g.
CC Parkinson's disease and Alzheimer's disease), cancer, leukaemia,
CC and tumours. The ACHE gene maps to human chromosome 7q22. The
CC present sequence is the coding sequence of the ACHE gene.
XX
SQ Sequence 1845 BP; 302 A; 590 C; 608 G; 345 T; 0 other;
Query Match 99.8%; Score 1722; DB 24; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 atgagggccccgcagtgctgctgcacacgccttccctggctcccaactccttcctc 60
|||||
Db 1 atgagggccccgcagtgctgctgcacacgccttccctggctcccaactccttcctc 60
Qy 61 ctccctggtcctcctgggtggaggagtgggggctgagggccggaggatgcagagctgctg 120
|||||
Db 61 ctccctggtcctcctgggtggaggagtgggggctgagggccggaggatgcagagctgctg 120
Qy 121 gtgacggtgcgtggggccggctgcggggcattcgcctgaagacccccggggccctgtc 180
|||||
Db 121 gtgacggtgcgtggggccggctgcggggcattcgcctgaagacccccggggccctgtc 180
Qy 181 tctgctttcctgggcatcccccttgcggagccacccatgggacccctcgtcttctgcc 240
|||||
Db 181 tctgctttcctgggcatcccccttgcggagccacccatgggacccctcgtcttctgcc 240
Qy 241 ccggagcccaagcagccttggtcaggggtggtgtagacgctacaacctccagagtgtctgc 300
|||||
Db 241 ccggagcccaagcagccttggtcaggggtggtgtagacgctacaacctccagagtgtctgc 300
Qy 301 taccaatatgtggacacccctataccccagggttttgaggccaccgagatgtggaacccaac 360
|||||
Db 301 taccaatatgtggacacccctataccccagggttttgaggccaccgagatgtggaacccaac 360
Qy 361 cgtgagctgagcgaggactgcctgtacctcaacgctgtggacaccataccccccggcctaca 420
|||||
Db 361 cgtgagctgagcgaggactgcctgtacctcaacgctgtggacaccataccccccggcctaca 420
Qy 421 tccccaccacctgtcctcgtctggatctatgggggtggcttctacagtggggcctctcc 480
|||||
Db 421 tccccaccacctgtcctcgtctggatctatgggggtggcttctacagtggggcctctcc 480

QY 481 ttggacgtgtacgatggccgcttcttggtacagggccgagaggactgtgctggtgtccatg 540
Db 481 ttggacgtgtacgatggccgcttcttggtacagggccgagaggactgtgctggtgtccatg 540
QY 541 aactaccgggtggagcccttttggtctcttgccctgcccctgcgcgggagccgagagggcccccgggc 600
Db 541 aactaccgggtggagcccttttggtctcttgccctgcccctgcgcgggagccgagagggcccccgggc 600
QY 601 aatgtgggtctctggatcagagagctggccctgcagtggtgcagggagaaactgtggcagcc 660
Db 601 aatgtgggtctctggatcagagagctggccctgcagtggtgcagggagaaactgtggcagcc 660
QY 661 ttcgggggtgacccgacatcagtgacgtgttttgggagagcgcgggacccctcggtg 720
Db 661 ttcgggggtgacccgacatcagtgacgtgttttgggagagcgcgggacccctcggtg 720
QY 721 ggcattgcacctgtgtccccgcccagccgggcccgtttccacagggcccgtgctgcagagc 780
Db 721 ggcattgcacctgtgtccccgcccagccgggcccgtttccacagggcccgtgctgcagagc 780
QY 781 ggtgcccccaatggacccttgggcccacggtgggcatgggagagggcccgtgcagggccacg 840
Db 781 ggtgcccccaatggacccttgggcccacggtgggcatgggagagggcccgtgcagggccacg 840
QY 841 cagctggcccaccttgtgggtgtctctccagggcgcactggtgggaatgacacagagctg 900
Db 841 cagctggcccaccttgtgggtgtctctccagggcgcactggtgggaatgacacagagctg 900
QY 901 gtacgtgaccttcgggacacgaccagcgcaggtctctggtgaaccacgaaatggcacgtgctg 960
Db 901 gtacgtgaccttcgggacacgaccagcgcaggtctctggtgaaccacgaaatggcacgtgctg 960
QY 961 cctcaagaagcgtcttccggttctctcgttgcctgtggtagatggagacttccctcagt 1020
Db 961 cctcaagaagcgtcttccggttctctcgttgcctgtggtagatggagacttccctcagt 1020
QY 1021 gacacccacagagccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
Db 1021 gacacccacagagccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
QY 1081 gtggtgaaggatgagggcctcgtattttctggttttacggggccccagggcttcagcaaaagac 1140
Db 1081 gtggtgaaggatgagggcctcgtattttctggttttacggggccccagggcttcagcaaaagac 1140
QY 1141 aacgagtcctcatcagccgggcccaggttctctgcccgggtgcgggtgcgggttccccag 1200
Db 1141 aacgagtcctcatcagccgggcccaggttctctgcccgggtgcgggtgcgggttccccag 1200
QY 1201 gtaagtgcacctggcagccgagggctgtggttcctgcattacacagactggctgcaccccgag 1260
Db 1201 gtaagtgcacctggcagccgagggcctgagcgtgtggtggcgaccacaatgtcgtgtgc 1260
QY 1261 gaccggcacgcctgagggagggccctgagcgtgtggtggcgaccacaatgtcgtgtgc 1320
Db 1261 gaccggcacgcctgagggagggcctgagcgtgtggtggcgaccacaatgtcgtgtgc 1320
QY 1321 cccgtggcccagctggctgggcgactggctggcccgaggtgcccgggtctacgcctacgtc 1380
Db 1321 cccgtggcccagctggctgggcgactggctggcccgaggtgcccgggtctacgcctacgtc 1380
QY 1381 tttgaacacccgtgcttccacgctctcctgcccctgtggatgggggtgcccacagggctac 1440
Db 1381 tttgaacacccgtgcttccacgctctcctgcccctgtggatgggggtgcccacagggctac 1440
QY 1441 gagatcgagttcatcttttgggatcccccttggaacccctctcgaaactacacggcagggag 1500
Db 1441 gagatcgagttcatcttttgggatcccccttggaacccctctcgaaactacacggcagggag 1500
QY 1501 aaaatcttcgccagcgactgatgcgatacttgggccaaactttgcccgcacaggggatccc 1560
Db 1501 aaaatcttcgccagcgactgatgcgatacttgggccaaactttgcccgcacaggggatccc 1560
QY 1561 aatgagcccccgagaccccaaggccccacaaatggcccccggtacacggcgggggtcagcag 1620

Db 1561 aatgagccccgagaccccaaggccccacaaatggccccgtacacggcgggggtcagcag 1620
QY 1621 tacgttagtctggacctgcccgtgcccgtggaggtgcggcgggggtgcgcgcccagggcctgc 1680
Db 1621 tacgttagtctggacctgcccgtgcccgtggaggtgcggcgggggtgcgcgcccagggcctgc 1680
QY 1681 gccttctggaacgcgttccctccccaaattgctcagcgccacc 1722
Db 1681 gccttctggaacgcgttccctccccaaattgctcagcgccacc 1722
RESULT 3
AAQ99002
ID AAQ99002 standard; DNA; 2256 BP.
XX AAQ99002;
AC AAQ99002;
XX 31-MAR-1996 (first entry)
DT Human acetylcholinesterase (AChE) gene.
XX
DE Acetylcholinesterase; acetyl cholinesterase; EC-3.1.1.7;
XX chromosome-7q22; acetylcholine-hydrolyzing enzyme; ss.
KW Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 160..2206
FT /*tag= a
XX
PN WO9523158-A1.
XX
PD 31-AUG-1995.
XX
PF 28-FEB-1995; 95WO-US02806.
XX
PR 09-JAN-1995; 95US-0370156.
PR 28-FEB-1994; 94US-0202755.
XX
PA (KOHN/) KOHN K I.
PA (YISS) YISSUM RES & DEV CO.
XX
PI Shani M, Soreq H, Zakut H;
XX
DR WPI; 1995-311499/40.
DR P-PSDB; AAR80726.
XX
PS Claim 3; Fig.1A; 55pp; English.
XX
CC This DNA sequence encoding human acetylcholinesterase is useful
CC for producing transgenic animals which express AChE. The
CC transgenic animals are in turn useful as an assay system for
CC determining the anti-ChE activity of organophosphates, carbamates,
CC anti-ChE drugs, plant glycoalkaloids and snake venoms. This gene
CC contains a promoter region, 6 exons (E1 to E6) and 4 introns (I1-
CC I4).
XX
SQ Sequence 2256 BP; 390 A; 757 C; 680 G; 429 T; 0 other;

Query Match 99.8%; Score 1722; DB 16; Length 2256;
Best local similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagggcccccgagtgctgtgcacacgccttccctggcttccccactccttctctc 60
Db 160 atgagggcccccgagtgctgtgcacacgccttccctggcttccccactccttctc 219

QY 61 ctccctctgctcctgggtgagagagtggggctgagggccgggaggatgcagagctgctg 120
|||||
Db 220 ctccctctgctcctgggtgagagagtggggctgagggccgggaggatgcagagctgctg 279
|||||
QY 121 gtgacggtgcgtggggccggtgcggggcatttcgctcctgaagacccccggggccctgtc 180
|||||
Db 280 gtgacggtgcgtggggccggtgcggggcatttcgctcctgaagacccccggggccctgtc 339
|||||
QY 181 tctgctttcctgggcatcccttttcgctcaggggtgtagacgtacacaccttccagagtgtctgc 240
|||||
Db 340 tctgctttcctgggcatcccttttcgctcaggggtgtagacgtacacaccttccagagtgtctgc 399
|||||
QY 241 ccggagcccaagcagccttggtcaggggtgtagacgtacacaccttccagagtgtctgc 300
|||||
Db 400 ccggagcccaagcagccttggtcaggggtgtagacgtacacaccttccagagtgtctgc 459
|||||
QY 301 taccaatatgtggacacctataccaggttttgagggcaaccgagatgtggaacccaac 360
|||||
Db 460 taccaatatgtggacacctataccaggttttgagggcaaccgagatgtggaacccaac 519
|||||
QY 361 cgtgagctgagcagggactgcctgtacctaactgtgtgacacaccataccccggcctaca 420
|||||
Db 520 cgtgagctgagcagggactgcctgtacctaactgtgtgacacaccataccccggcctaca 579
|||||
QY 421 tccccaccctgtcctcgtctggatctatgggggtggttctacagtggggacctctcc 480
|||||
Db 580 tccccaccctgtcctcgtctggatctatgggggtggttctacagtggggacctctcc 639
|||||
QY 481 ttggacgtgtacgatggccgcttcttggtacaggcccgagagactgtgctggtgtccatg 540
|||||
Db 640 ttggacgtgtacgatggccgcttcttggtacaggcccgagagactgtgctggtgtccatg 699
|||||
QY 541 aactaccgggtgggagcctttggcttccctggccctgcggggagccgagagccccgggc 600
|||||
Db 700 aactaccgggtgggagcctttggcttccctggccctgcggggagccgagagccccgggc 759
|||||
QY 601 aatgtgggtcctcctggatcagaggtcggccctgcagtggtgtaggagaaactgtggaacc 660
|||||
Db 760 aatgtgggtcctcctggatcagaggtcggccctgcagtggtgtaggagaaactgtggaacc 819
|||||
QY 661 ttccgggggtgaccccgacatcagtgacgtgctgttggggagagcgcgggacctcgggtg 720
|||||
Db 820 ttccgggggtgaccccgacatcagtgacgtgctgttggggagagcgcgggacctcgggtg 879
|||||
QY 721 ggcattgcacctgctgtcccgcccgccagccgggacctgttccacagggccgtgtgcagagc 780
|||||
Db 880 ggcattgcacctgctgtcccgcccgccagccgggacctgttccacagggccgtgtgcagagc 939
|||||
QY 781 ggtgcccccaaatggaccctggggccacaggtgggcatgggagagggcccgtcgagggccaag 840
|||||
Db 940 ggtgcccccaaatggaccctggggccaaggtgggcatgggagagggcccgtcgagggccaag 999
|||||
QY 841 cagctggccccacctgtgtgggtctcctccaggcggcactggtgggaatgacacagagctg 900
|||||
Db 1000 cagctggccccacctgtgtgggtctcctccaggcggcactggtgggaatgacacagagctg 1059
|||||
QY 901 gtgacctgccttcggacacgacctgagcaggtcctcctcgtgcctgtggtagatggagacttccag 960
|||||
Db 1060 gtgacctgccttcggacacgacctgagcaggtcctcctggtgaaccacgaatggcacgtgctg 1119
|||||
QY 961 cctcaagaaagcgtcttcgggttctcctcgtgcctgtggttagatggagacttccag 1020
|||||
Db 1120 cctcaagaaagcgtcttcgggttctcctcgtgcctgtggttagatggagacttccag 1179
|||||
QY 1021 gacacccagaggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
|||||
Db 1180 gacacccagaggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1239
|||||
QY 1081 gtggtgaaggatgagggctcgtattttctggtttacggggccccaggcttcagcaaaagac 1140
|||||
Db 1240 gtggtgaaggatgagggctcgtattttctggtttacggggccccaggcttcagcaaaagac 1299
|||||
QY 1141 aacgagctctcatcagccggcgaggttccctggccggggtgcgggtcggggtcccccag 1200
|||||

Db 1300 aacgagctctcatcagccggcgagttcctgcccgggtgcgggttccccag 1359
|||||
QY 1201 gtaagtgaacctggcagccgaggtggtgctcctgcattacagagctggctgcacccag 1260
|||||
Db 1360 gtaagtgaacctggcagccgaggtggtcctgcattacagagctggctgcacccag 1419
|||||
QY 1261 gaccggcacgcctgaggggagggccctgagcgtggtggtgggagaccacaatgctgtgtc 1320
|||||
Db 1420 gaccggcacgcctgaggggagggccctgagcgtggtggtgggagaccacaatgctgtgtc 1479
|||||
QY 1321 ccgctggccccagctggctggggagactggctgccagggtgccccgggtctacgcctacgtc 1380
|||||
Db 1480 ccgctggccccagctggctggggagactggctgccagggtgccccgggtctacgcctacgtc 1539
|||||
QY 1381 ttggaacacccgtgtcttccacgtctcctggtgccccctgtggtggtgggtgccccacggctac 1440
|||||
Db 1540 ttggaacacccgtgtcttccacgtctcctggtgccccctgtggtggtgggtgccccacggctac 1599
|||||
QY 1441 gagtcgagttcatcttttgggataccccctggacccctctcgaactacacggcagaggag 1500
|||||
Db 1600 gagtcgagttcatcttttgggataccccctggacccctctcgaactacacggcagaggag 1659
|||||
QY 1501 aaaatcttcgcccagcactgatgcatactgggccaactttgcccgcacaggggatccc 1560
|||||
Db 1660 aaaatcttcgcccagcactgatgcatactgggccaactttgcccgcacaggggatccc 1719
|||||
QY 1561 aatgagccccgagacccccaaaggcccaacaatggcccccgctacacggcggtcagcag 1620
|||||
Db 1720 aatgagccccgagacccccaaaggcccaacaatggcccccgctacacggcggtcagcag 1779
|||||
QY 1621 tacgttagtctggacctgcggcgctggaggtgcggcggtgctgcgcgccagggcctgc 1680
|||||
Db 1780 tacgttagtctggacctgcggcgctggaggtgcggcggtgctgcgcgccagggcctgc 1839
|||||
QY 1681 gccttctggaaccgcttctcctcccaaaattgctcagcgccacc 1722
|||||
Db 1840 gccttctggaaccgcttctcctcccaaaattgctcagcgccacc 1881
|||||

RESULT 4
AAS17547

ID AAS17547 standard; DNA; 5767 BP.

XX

AC AAS17547;

XX

DT 25-FEB-2002 (first entry)

XX

DE Plasmid pTM034.

XX

KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
transgenic plant; acetylcholinesterase poisoning; chemical warfare;
muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigmine bromide; pTM034;
KW circular; ds.

XX

OS Synthetic.

XX

PN WO200171014-A2.

XX

PD 27-SEP-2001.

XX

PF 16-MAR-2001; 2001WO-US08468.

XX

PR 17-MAR-2000; 2000US-190440P.

XX

PA (MORT/) MOR T.

PA (SORE/) SOREQ H.

PA (ARNT/) ARNTZEN C.

PA (MASO/) MASON H.

PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.

XX

PI Mor T, Soreq H, Arntzen C, Mason H;

|||||
Db 2398 aatgagccccgagaccccaaggccccacaatggcccccgctacacggcgggggtcagcag 2457
QY 1621 tacgttagcttgacactgcggccgctggaggtgcggcgggggtgcgcgcagccagcctgc 1680
Db 2458 tacgttagcttgacactgcggccgctggaggtgcggcgggggtgcgcgcagccagcctgc 2517
QY 1681 gccttctggaaccgcttctctcccccataattgctcagcgccacctga 1725
Db 2518 gccttctggaaccgcttctctcccccataattgctcagcgctacctga 2562
RESULT 5
AAS17548
ID AAS17548 standard; DNA; 14446 BP.
XX
AC AAS17548;
XX
DT 25-FEB-2002 (first entry)
XX
DE Plasmid pTM036.
XX
KW Human; acetylcholinesterase; AChE; cyclic; antidote; pesticide;
KW transgenic plant; acetylcholinesterase poisoning; chemical warfare;
KW muscarinic receptor antagonist; atropine; oxime; anti-cholinergic;
KW organophosphate(OP)-modified AChE; pyridostigamine bromide; pTM036;
KW circular; ds.
XX
OS Synthetic.
XX
PN WO200171014-A2.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-US08468.
XX
PR 17-MAR-2000; 2000US-190440P.
XX
PA (MORT/) MOR T.
PA (SORE/) SOREQ H.
PA (ARNT/) ARNTZEN C.
PA (MASO/) MASON H.
PA (BOYC-) BOYCE THOMPSON INST PLANT RES INC.
XX
PI Mor T, SoREQ H, Arntzen C, Mason H;
XX
DR WPI; 2002-055120/07.
XX
PT Production of a transgenic plant which contains a polynucleotide that
PT encodes a human acetylcholinesterase which upon purification is
PT effective against acetylcholinesterase poisoning -
XX
PS Claim 11; Page 32-41; 42pp; English.
XX
CC The invention relates to a method of producing a transgenic plant which
CC contains a polynucleotide that encodes human acetylcholinesterase (AChE)
CC which upon purification is effective against acetylcholinesterase
CC poisoning. The method is used for treating a victim of
CC acetylcholinesterase poisoning by administering a therapeutic amount of
CC a physiologically active human acetylcholinesterase expressed in plant
CC tissue. The extensive use of anticholinesterase pesticides with
CC concurrent accidental poisoning, the threat of chemical warfare and
CC environmental concerns demand the development of effective, inexpensive
CC and stage countermeasures and bioremediation solutions. Prior art methods
CC for treating AChE poisoning have used the muscarinic receptor antagonist
CC atropine and oximes to reactivate the organophosphate(OP)-modified AChE.
CC The reversible carbamate, pyridostigamine bromide has also been used as a
CC prophylactic. However, these conventional treatments have limited
CC effectiveness and serious short and long-term side effects and may result
CC in significant performance deficits and even permanent brain damage. This
CC invention permits the utilisation of cholinesterases to counter-act the
CC toxic effects of anti-cholinergic agents. Using transgenic plants for the
CC production of the enzymes is cost effective and the product is stable

CC so that the injected enzymes have the advantage of having a long
CC half-life. The transgenic form of the enzymes are also easy to purify.
CC The present sequence is plasmid pTM036, the pGPTVkan derivative construct
CC used in the generation of transgenic tomato plants that constitutively
CC express human AChE.
XX
SQ Sequence 14446 BP; 3231 A; 3831 C; 4046 G; 3042 T; 296 other;
Query Match 99.6%; Score 1718.6; DB 24; Length 14446;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1721; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 atgagggccccgcagtgctctgtgcacacgccttccctggcttccccactccttctctc 60
Db ||| |||||
Db 142 atggctccccgcagtgctctgtgcacacgccttccctggcttccccactccttctctc 201
QY 61 ctctcttggtcctcctgggtggaggagtggtgggctgagggccgggaggatgcagagctgtg 120
Db |||||
Db 202 ctctctggctcctcctgggtggaggagtggtgggctgagggccgggaggatgcagagctgtg 261
QY 121 gtgacggtgcgtggggcggtgcgtggggcattcgctcctgaagacccccggggcctgtgc 180
Db |||||
Db 262 gtgacggtgcgtggggcggtgcgtggggcattcgctcctgaagacccccggggcctgtgc 321
QY 181 tctgcttctcctgggcatcccttctgcggagccaccctatgggacccctgcgtcttctgcga 240
Db |||||
Db 322 tctgcttctcctgggcatcccttctgcggagccaccctatgggacccctgcgtcttctgcga 381
QY 241 ccggagcccaagcagccttgggtcagggtggtgtagagcctacaaccttccagagtgctgtgc 300
Db |||||
Db 382 ccggagcccaagcagccttgggtcagggtggtgtagagcctacaaccttccagagtgctgtgc 441
QY 301 taccataatgtggacacccctatataccagggttttgagggcacccagatgtggaaccccaac 360
Db |||||
Db 442 taccataatgtggacacccctatataccagggttttgagggcacccagatgtggaaccccaac 501
QY 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccccggcctaca 420
Db |||||
Db 502 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccccggcctaca 561
QY 421 tccccacccctgtcctcgtctggatctatggtgggtggttctctacagtggtggtcctcc 480
Db |||||
Db 562 tccccacccctgtcctcgtctggatctatggtgggtggttctctacagtggtggtcctcc 621
QY 481 ttggacgtgtacgatggcgccttcttggtacagggccgagagactgtgctggtgtccatg 540
Db |||||
Db 622 ttggacgtgtacgatggcgccttcttggtacagggcagagagactgtgctggtgtccatg 681
QY 541 aactaccgggtggagccttggcttctcctggcctcgcggggagccgagagcccggggc 600
Db |||||
Db 682 aactaccgggtggagccttggcttctcctggcctgcggggagccgagagcccggggc 741
QY 601 aatgtgggtcctcctggatcagaggctggccctgcagtggtgcaggagaacgtggcagcc 660
Db |||||
Db 742 aatgtgggtcctcctggatcagaggctggccctgcagtggtgcaggagaacgtggcagcc 801
QY 661 ttccggggtgacccgacatcagtgcgtgtttgggagagcgcgggagccgcctcgtgtg 720
Db |||||
Db 802 ttccggggtgacccgacatcagtgcgtgtttgggagagcgcgggagccgcctcgtgtg 861
QY 721 ggcattgcacctgcttcccgcgccagccggggcctgttccacagggccgtgctgcagagc 780
Db |||||
Db 862 ggcattgcacctgcttcccgcgccagccggggcctgttccacagggccgtgctgcagagc 921
QY 781 ggtgcccccaatggacctggggccacggtgggcatgggagagggcccgctgcagggccacg 840
Db |||||
Db 922 ggtgcccccaatggacctggggccacggtgggcatgggagagggcccgctgcagggccacg 981
QY 841 cagctggcccaccttgtgggctgtcctccagggcggaactggtgggaatgacacagagctg 900
Db |||||
Db 982 cagctggcccaccttgtgggctgtcctccagggcggaactggtgggaatgacacagagctg 1041

QY 901 gtagcctgccttcggacacgaccagcgaggctcctggtgaaccacgaatggcacgtgctg 960
Dbb 1042 gtagcctgccttcggacacgaccagcgaggctcctggtgaaccacgaatggcacgtgctg 1101
QY 961 cctcaagaagaagcgtcttcocggtctccttcctgctgctgtagatggagacttccctcagt 1020
Dbb 1102 cctcaagaagaagcgtcttcocggtctccttcctgctgctgtagatggagacttccctcagt 1161
QY 1021 gacacccagaggccctcatcaacgcgaggagacttccacggccctgcaggtgctggtgggt 1080
Dbb 1162 gacacccagaggccctcatcaacgcgaggagacttccacggccctgcaggtgctggtgggt 1221
QY 1081 gtggtgaaggatgagggctcgtattttctggttacggggcccccaggtctcagcaaaagac 1140
Dbb 1222 gtggtgaaggatgagggctcgtattttctggttacggggcccccaggtctcagcaaaagac 1281
QY 1141 aacgagtcctcatcagccgggcccaggttccttgccggggtgcgggttcctcccaag 1200
Dbb 1282 aacgagtcctcatcagccgggcccaggttccttgccggggtgcgggttcctcccaag 1341
QY 1201 gtaagtgaacctggcagccgaggtggttcctgcattacacagactggctgcacatcccgag 1260
Dbb 1342 gtaagtgaacctggcagccgaggtggttcctgcattacacagactggctgcacatcccgag 1401
QY 1261 gacccggcacgcctgagggagggccctgagcgatgtggtggcgaccacaaatgtcgtgtgc 1320
Dbb 1402 gaccggcacgcctgagggagggccctgagcgatgtggtggcgaccacaaatgtcgtgtgc 1461
QY 1321 ccggtggccagctggctggcgagctggctgcccagggtgcggccgggtctacgcctacgtc 1380
Dbb 1462 ccggtggccagctggctggcgagctggctgcccagggtgcggccgggtctacgcctacgtc 1521
QY 1381 ttgaaacacgtgtcttccacgctctcctgcccctgtggatgggggtgccccacgggtac 1440
Dbb 1522 ttgaaacacgtgtcttccacgctctcctgcccctgtggatgggggtgccccacgggtac 1581
QY 1441 gagatcgagttcatctttgggatccccctggaccccctctcgaaactacacggcgagaggag 1500
Dbb 1582 gagatcgagttcatctttgggatccccctggaccccctctcgaaactacacggcgagaggag 1641
QY 1501 aaaaatcttcgcccagcgactgatcgatactgaggcccaactttgcccgcacaggggatccc 1560
Dbb 1642 aaaaatcttcgcccagcgactgatcgatactgaggcccaactttgcccgcacaggggatccc 1701
QY 1561 aatgagcccccgagaccccccaaggccccacaatggcccccgatcacagggggtcagcag 1620
Dbb 1702 aatgagcccccgagaccccccaaggccccacaatggcccccgatcacagggggtcagcag 1761
QY 1621 tacgttagtctggacctgcgcgcgcgtggaggtgcggcgggggtgcgcgccccaggtgc 1680
Dbb 1762 tacgttagtctggacctgcgcgcgcgtggaggtgcggcgggggtgcgcgccccaggtgc 1821
QY 1681 gccttctggaaccgcttctcccccataattgctcagcgccacctga 1725
Dbb 1822 gccttctggaaccgcttctcccccataattgctcagcgctacctga 1866

RESULT 6
AAQ05998
ID AAQ05998 standard; DNA; 2253 BP.
XX
AC AAQ05998;
XX
DT 16-JAN-1991 (first entry)
XX
DE Sequence encoding human acetylcholinesterase (hAChE).
XX
KW Organophosphorous poisoning; OP; cancer; leukaemia;
KW megakaryocytopoiesis; ovarian cancer; ds.
XX
OS Homo sapiens.
XX
PN EP388906-A.

XX 26-SEP-1990.
XX
XX 20-MAR-1990; 90EP-0105274.
XX
PR 21-MAR-1989; 89IL-0089703.
XX
PA (YISS) YISSUM RES DEV CO.
XX
XX Soreq H, Zakut H;
PI
XX WPI; 1990-291865/39.
DR P-PSDB; AAR06989.
XX
PT Human acetylcholinesterase DNA and prodn. of recombinant hAChE -
PT for treatment of organo-phosphorous poisoning and diagnosis of haemo-
PT cytopoeitic disorders and ovarian carcinomas.
XX
PS Claim 5; Page 25; 47pp; English.
XX
CC Gene product is useful as an active pharmaceutical component for the
CC prophylaxis and treatment of organophosphorous poisoning, and post-
CC surgical apnea due to succinylcholine administration.
CC cDNA probe to the sequence may be used in diagnosis of various
CC leukaemias, abnormal megakaryocytopoiesis and ovarian carcinomas.
XX
SQ Sequence 2253 BP; 390 A; 740 C; 694 G; 429 T; 0 other;

Query Match 97.7%; Score 1684.8; DB 11; Length 2253;
Best Local Similarity 99.4%; Pred. NO. 0;
Matches 1712; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 1 atgagggcccccgcagtgctgtctgcacacgacctccctggcttccccactccttctcctc 60
Dbb 160 atgagggcccccgcagtgctgtctgcacacgacctccctggcttccccactccttctcctc 219
QY 61 ctctctggtcctcctgggtggagagtggtgggctgagggcgaggagatgcagagctgctg 120
Dbb 220 ctctctggtcctcctgggtggagagtggtgggctgagggcgaggagatgcagagctgctg 279
QY 121 gtacggtgcgtggggccggtgcgggctgcattgcctgaagacccccggggccctgtc 180
Dbb 280 gtacggtgcgtggggccggtgcggg--attgcctgaaga-ccccggggggccctgtc 336
QY 181 tctgctttcctgggcatccccctttgcggagaccacccatgggacccccgtctttctgccca 240
Dbb 337 tctgctttcctgggcatccccctttgcggagaccacccatgggacccccgtctttctgccca 396
QY 241 cgggagcccaagcagccttggtcagggggtggtagacgctacaaccttccagagtgctgc 300
Dbb 397 cgggagcccaagcagccttggtcagggggtggtagacgctacaaccttccagagtgctgc 456
QY 301 taccataatgtggacacccctataccccaggttttggggcgacccgagatgtggaaccccaac 360
Dbb 457 taccataatgtggacacccctataccccaggttttggggcgacccgagatgtggaaccccaac 516
QY 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccggtctaca 420
Dbb 517 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccggtctaca 576
QY 421 tccccacccccctgtcctcgtctggatctatgggggtggtcttctacagtggggacctctcc 480
Dbb 577 tccccacccccctgtcctcgtctggatctatgggggtggtcttctacagtggggacctctcc 636
QY 481 ttggacgtgtacgatggccgcttcttgggtacaggccgagaggactgtgtggtgtccatg 540
Dbb 637 ttggacgtgtacgatggccgcttcttgggtacaggccgagaggactgtgtggtgtccatg 696
QY 541 aactaccgggtgggagcctttggcttccctggccctgcccggggagccgagagggccggggc 600
Dbb 697 aagtaccgggtgggagcctttggcttccctggccctgcccggggagccgagagggccggggc 756

Qy 601 aatgtgggtctcttgatcagagggctggccctgcagtggtgtcagagaaacgtggcagcc 660
Db 757 aatgtgggtctcttgatcagagggctggccctgcagtggtgtcagagaaacgtggcagcc 816
Qy 661 ttcgggggtgacccgacatcagtgacgtgtttggggagagcgcgggagccgctcggtg 720
Db 817 ttcgggggtgacccgacatcagtgacgtgtttggggagagcgcgggagcccgctcggtg 876
Qy 721 ggcataccactgctgtccccgcagcccggggcctgttccacacagggccgctgctgcagagc 780
Db 877 ggcataccactgctgtccccgcagcccggggcctgttccacacagggccgctgctgcagagc 936
Qy 781 ggtgcccccaatggaccctgggccacgggtgggcacatgggagagggcccgctgcagggccacg 840
Db 937 ggtgcccccaatggaccctgggccacgggtgggcacatgggagagggcccgctgcagggccac 996
Qy 841 cagctggcccaacctgtgtggctgtctccagcgggcactggtgggaatgacacagagctg 900
Db 997 cagctggcccaacctgtgtggctgtctccagcgggcactggtgggaatgacacagagctg 1056
Qy 901 gtagcctgccttcggaacgacgagcgaggtcctggtgaaccacgaatggcacgtgctg 960
Db 1057 gtagcctgccttcggaacgacgagcgaggtcctggtgaaccacgaatggcacgtgctg 1116
Qy 961 cctcaagaaagcgtcttcgggttctccttcgtgcctgtggtagatggagacttcctcagt 1020
Db 1117 cctcaagaaagcgtcttcgggttctccttcgtgcctgtggtagatggagacttcctcagt 1176
Qy 1021 gacaccccagagggccctcatcaacggggagacattccacggcctgcaggtgctgggggt 1080
Db 1177 gacaccccagagggccctcatcaacggggagacattccacggcctgcaggtgctgggggt 1236
Qy 1081 gtggtgaaggatgagggtcgtctatttctgtttacggggcccgaggttcagcaaagac 1140
Db 1237 gtggtgaaggatgagggtcgtctatttctgtttacggggcccgaggttcagcaaagac 1296
Qy 1141 aacgagttctcatcagccgggcccaggttccctggtggggtgcgggtcgggttccccag 1200
Db 1297 aacgagttctcatcagccgggcccaggttccctggtggggtgcgggtcgggttccccag 1356
Qy 1201 gtaagtgaacctggcagccgaggtgctgtggtcctgcattacacagactggctgcaccccag 1260
Db 1357 gtaagtgaacctggcagccgaggtgctgtggtcctgcattacacagactggctgcaccccag 1416
Qy 1261 gacccggcacgcctgagggagggccctgagcgatgtggtgggcgacacaaatgctgtgc 1320
Db 1417 gacccggcacgcctgagggagggccctgagcgatgtggtgggcgacacaaatgctgtgc 1476
Qy 1321 cccgtggcccagctggctggggcagctggctggccaggggtgcggggtctacgcctacgtc 1380
Db 1477 cccgtggcccagctggctggggcagctggctggccaggggtgcggggtctacgcctacgtc 1536
Qy 1381 tttgaacacacgtgcttccacgctctcctggccccctgtggtggtgggtgccccacggctac 1440
Db 1537 tttgaacacacgtgcttccacgctctcctggccccctgtggtggtgggtgccccacggctac 1596
Qy 1441 gagatcgagttcatctttgggatccccctggacccctctcgaaactacacggcagaggag 1500
Db 1597 gagatcgagttcatctttgggatccccctggacccctctcgaaactacacggcagaggag 1656
Qy 1501 aaaatcttgcggcagcgactgatgcatactgggccaactttgccgcgacaggggatccc 1560
Db 1657 aaaatcttgcggcagcgactgatgcatactgggccaactttgccgcgacaggggatccc 1716
Qy 1561 aatgagccccgagagaccccaaggcccccaaatggcccccgtagacggggggtctcagcag 1620
Db 1717 aatgagccccgagagaccccaaggcccccaaatggcccccgtagacggggggtctcagcag 1776
Qy 1621 tacgttagtctggacctgcggccgctgaggtgcggcgggggtgcgcgccccaggcctgc 1680
Db 1777 tacgttagtctggacctgcggccgctgaggtgcggcgggggtgcgcgccccaggcctgc 1836
Qy 1681 gccttctggaaccgcttctctccccaaattgctcagcgccacc 1722

Db 1837 gccttctggaaccgcttctctccccaaattgctcagcgccacc 1878
RESULT 7
AAQ05999
ID AAQ05999 standard; DNA; 1800 BP.
XX
AC AAQ05999;
XX
DT 16-JAN-1991 (first entry)
XX
DE Sequence encoding foetal human acetylcholinesterase (hAChE).
XX
KW Organophosphorous poisoning; OP; cancer; leukaemia;
megakaryocytopoiesis; ovarian cancer; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1500
FT /*tag= a
FT 3'UTR 1501..1800
FT /*tag= b
XX
PN EP388906-A.
XX
PD 26-SEP-1990.
XX
PF 20-MAR-1990; 90EP-0105274.
XX
PR 21-MAR-1989; 89IL-0089703.
XX
PA (YISS) YISSUM RES DEV CO.
XX
PI Soreq H, Zakut H;
XX
DR WPI; 1990-291865/39.
DR P-PSDB; AAR06990.
XX
PT Human acetyl:cholinesterase DNA and prodn. of recombinant hAChE -
PT for treatment of organo-phosphorous poisoning and diagnosis of haemo-
PT cytopoeitic disorders and ovarian carcinomas.
XX
PS Disclosure; Fig 1c; 47pp; English.
XX
CC hAChE is useful as an active pharmaceutical component for the
CC prophylaxis and treatment of organophosphorous poisoning, and
CC post-surgical apnea due to succinylcholine administration.
CC cDNA probe to the sequence may be used in diagnosis of various
CC leukaemias, abnormal megakaryocytopoiesis and ovarian carcinomas.
CC Bases given as N are obscured in the patent specification.
XX
SQ Sequence 1800 BP; 330 A; 602 C; 539 G; 326 T; 3 other;

Query Match 76.2%; Score 1313.6; DB 11; Length 1800;
Best Local Similarity 98.4%; Pred. No. 5.6e-268;
Matches 1358; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

Qy 346 atgtggaaccccaaccgtgagctgagcgaggactgcctgtacctcaacgtgtggacacca 405
Db 1 atgtggaaccccaaccgtgagctgagcgaggactgcctgtacctcaacgtgtggacacca 60
Qy 406 taccocccggcctacatccccacccctgtcctcgtctggtatctatgggggtggttctac 465
Db 61 taccocccggcctacatccccacccctgtcctcgtctggtatctatgggggtggttctac 120
Qy 466 agtggggcctccttggacgtgtacggtgacggtgcccgttcttggtaacggccgagaggact 525
Db 121 agtggggcctccttggacgtgtacggtgcccgttcttggtaacggcccgagaccact 180
Qy 526 gtgctggtgtccatgaactaccgggtgggagcctttggcttctcctggccctgcccgggagc 585

Qy 121 gtgacgtgctggtgggcccggctgcggggcattcgccctgaagaccgccgggggcccctgtc 180
Db 121 gtgaggttcgagggggccagctgaggggcatccgcctgaaggccctggagggccagtc 180
Qy 181 tctgctttcctgggcatccccctttfaggagccaccatgggaccccgctgcgtttctgcc 240
Db 181 tcagcttttctgggcatccccctttgcagagccacctgtgggtcacgttagatttatgcc 240
Qy 241 ccggagcccaagcagcccttggtcaggggtggtagacgctacaaccttccagagtgtctgc 300
Db 241 ccagagcccaagcggccctggtcagagagtgttggtgctaccaccttccaaaaatgtctgc 300
Qy 301 taccataatgtggacaccctataaccagggttttgagggcaccgagatgtggaaccccaac 360
Db 301 taccagtacgtggacaccctgtaccctgggtttgagggtactgagatgtggaaccccaac 360
Qy 361 cgtgagctgagcgaggactgcctgtactcaacgtgtggacaccatacccccgccctaca 420
Db 361 cgagagttagtgagactgcctgtatcttaatgtgtggacaccataccccagacctgtc 420
Qy 421 tccccaccctgtcctcgtctggtatctatggggttggtcttctacagtgggcccctccc 480
Db 421 tctcccaacctgtcctcatctggtatctatggggttggtttctacagcgagcgccctcc 480
Qy 481 ttggacgtgtacgatggccgcttcttggtacagccgagagagactgtgctggtccatg 540
Db 481 ttggatgtgtatgacggccgcttcttggtccaggttgaggagctgtgtgtatctatg 540
Qy 541 aactacccgggtgggagccctttggcttctcctgcccctgcggggagccgagagggcccgggc 600
Db 541 aactacccagtgggaaacctttggcttcttgcccctaccagggaagcagagaagcccctggc 600
Qy 601 aatgtgggtctcctggatcacagaggtcgccctgcagtggtgacgagaaacgtggcagcc 660
Db 601 aatgtaggtctgtggtacaaaggcttgcccttgcaatgggtgcaagaaaaatattgcagcc 660
Qy 661 ttccgggggtgacccgacatcagtgacgtgtttggggagagcgggagccgctcggtg 720
Db 661 tttggggggaccccgatgtcagtgactctgtttggggagagtgcgggtgcagccctcogtg 720
Qy 721 ggcattgcactgtgtccccgcccagcggggcctgttccacagggcctgtgcagagc 780
Db 721 ggcattgcacatactgtccctgcccagcaggagcctcttccacagggctgtcctccagagt 780
Qy 781 ggtgcccccaatggaccctgggccacgggtgggcatgggagagggcccgtgcgaggccacg 840
Db 781 ggcacacccaatgggccctgggccactgtgagtgtgtagagagggccagggcgccaca 840
Qy 841 cagctggcccaaccttgtgggctgtcctccagggcgcaactggtggaatgacacagagctg 900
Db 841 ctgctggcccgccttgtgggctgtccccccaggtggcgctggtggcaatgacaccgagctg 900
Qy 901 gtacccctgccttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
Db 901 atagcctgcttgaggacaaggcccgcctcaggacctggtggaccacgagtggcacgtcctg 960
Qy 961 cctcaagaaagcgtcttccggttctcctcgtgcctgtggttagatggagacttccctcagt 1020
Db 961 cctcaagaaagtatcttccgatttctcctcgtgcctgtggttagacggggacttccctcagt 1020
Qy 1021 gacacccccagaggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
Db 1021 gacacacgggagggctctcatcaatactaggagatttccaagacctgcaggtgctggtgggt 1080
Qy 1081 gtgggtgaaggatgagggctcgttattttctggtttacggggccccagggcttcagcaaaagac 1140
Db 1081 gtgggtgaaggacgaggggtcctacttcttctggtttacgggggttcccaaggcttcagcaaaagac 1140
Qy 1141 aacgagtctctcatcagccggcccgagttcctggccggggtgcgggtcgggttccccag 1200
Db 1141 aatgaatctctcatcagccgggcccagttcctggctgggtgcgggagcgtgtaccacca 1200
Qy 1201 gtaagtgcctggcagccgagggtgtggtcctgcattacacagactggctgcatccccgag 1260

Db 1201 gcaagtgaacctggcgccgaggtgtggtcctgcattacacagactggtgcacacctgag 1260
Qy 1261 gaccgggcacgcctgagggagggccctgagcgtgtggtggcgacacaaatgtcgtgtgc 1320
Db 1261 gaccctactcacctgagagatgccatgagtgcagtgtgtaggcgacacaaactgtgtgc 1320
Qy 1321 cccgtggcccagctggctggcgactggctgcccaggggtgcccggttctacgcctacgtc 1380
Db 1321 cctgtggcccagctggctggcgactggctgcccagggggcccgggtctatgcctacatc 1380
Qy 1381 ttggaacaccgtgcttccacgctctcctgcccctgttgatgggggtgcccacaggctac 1440
Db 1381 ttggaacaccgtgcttccacactgacttggtgcccctctggtatgggggtgcccctat 1440
Qy 1441 gagatcgagttcatcttttgggatccccctggaccccctctcgaaaactacacggcgagag 1500
Db 1441 gaaatcgagttcatcttttgggtcccccttgatccccctcgctgaactacacggagag 1500
Qy 1501 aaaatcttcgcccagcagctgatgcgatacttggtgcccctgtacacggcggtgctcagcag 1560
Db 1501 aggatcttctcagcgacttatgaaatactggaccaaatttgcgcgcacaggggacccc 1560
Qy 1561 aatgagcccgcagaccccccaaggcccccaaatggccccctgtacacggcggtgctcagcag 1620
Db 1561 aatgacccctcgagactccaaatctccacagtgggccaccgtacacactgcgcgcagcaa 1620
Qy 1621 tacgttagtctggacctgcggccgctggaggtgcggcggtgcggggctgcgcgcacaggcctgc 1680
Db 1621 tatgtgagcctgaacctgaagcccttagaggtgcggcggtgactgcgcgcacagacctgc 1680
Qy 1681 gccttctggaaccgcttctcccccaaatgtctcagcgccacc 1722
Db 1681 gccttctggaatcgcttctcccccaaatgtctcagcgccacc 1722

RESULT 9
AAS17492
ID AAS17492 standard; DNA; 12113 BP.
XX
AC AAS17492;
XX
DT 26-FEB-2002 (first entry)
XX
DE Human acetylcholinesterase, ACHE, gene.
XX
KW Human; ds; ACHE; acetylcholinesterase; polymorphic variant; haplotyping;
KW genotyping; neurological disease; Parkinson's disease;
KW Alzheimer's disease; cancer; leukaemia; tumour; chromosome 7q22.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace (1950,C)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (2237,T)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace (4044,T)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT CDS 4095..8113
FT /*tag= d
FT /product= "ACHE"
FT exon 4095..5162
FT /*tag= e
FT /number= 2
FT intron 5163..5508
FT /*tag= f
FT /number= 2
FT variation replace (4130,T)
FT /*tag= g

FT /standard_name= "Single nucleotide polymorphism"
FT replace (4195,A)
FT /*tag= h
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4277,C)
FT /*tag= i
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4932,G)
FT /*tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT replace (4967,T)
FT /*tag= k
FT /standard_name= "Single nucleotide polymorphism"
FT replace (5126,A)
FT /*tag= l
FT /standard_name= "Single nucleotide polymorphism"
FT replace (5151,A)
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism"
FT 5509..5993
FT /*tag= n
FT /number= 3
FT 5994..6988
FT /*tag= o
FT /number= 3
FT replace (5871,T)
FT /*tag= p
FT /standard_name= "Single nucleotide polymorphism"
FT replace (6831,A)
FT /*tag= q
FT /standard_name= "Single nucleotide polymorphism"
FT 6989..7158
FT /*tag= r
FT /number= 4
FT 7159..7991
FT /*tag= s
FT /number= 4
FT 7992..8113
FT /*tag= t
FT /number= 5
FT replace (8114,A)
FT /*tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT replace (8200,A)
FT /*tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT replace (8227,A)
FT /*tag= w
FT /standard_name= "Single nucleotide polymorphism"
FT replace (8425,G)
FT /*tag= x
FT /standard_name= "Single nucleotide polymorphism"

WO200179219-A2.

25-OCT-2001.

11-APR-2001; 2001WO-US11853.

14-APR-2000; 2000US-197173P.

(GENA-) GENAISSANCE PHARM INC.
(KAZE/) KAZEMI A.

Bentivegna SC, Chew A, Choi JY, Koshy B;

WPI; 2002-055248/07.

P-PSDB; AAU11231.

PT New polymorphic variants comprising acetylcholinesterase (ACHE)
PT isogene, useful in expressing ACHE protein for use in screening for
PT candidate drugs to treat diseases related to ACHE activity, e.g.
PT neurological diseases or cancer -

XX
PS Claim 21; Fig 1; 79pp; English.
XX
CC The invention relates to a polynucleotide comprising a polymorphic
CC variant of an acetylcholinesterase (ACHE) gene or fragment, protein or
CC complement, the variant comprising an ACHE isogene defined by a haplotype
CC selected from haplotypes 1-20 listed in the specification. Also included
CC are methods for haplotyping and genotyping the ACHE gene of an
CC individual, a method for predicting a haplotype pair for the ACHE gene of
CC an individual, a method for identifying an association between a trait
CC and at least one haplotype or haplotype pair of ACHE gene, recombinant
CC nonhuman organisms transformed or transfected with the polynucleotide
CC where the organism expresses ACHE protein encoded by the first
CC nucleotide sequence or encoded by the polymorphic variant sequence,
CC an isolated antibody specific for and immunoreactive with ACHE,
CC a method of screening for drugs targeting the polypeptide contacting ACHE
CC polymorphic variant with a candidate agent and assaying for binding
CC activity, a computer system for storing and analysing polymorphism data
CC for ACHE gene and a genome anthology for ACHE gene which comprises ACHE
CC isogenes defined by haplotypes 1-20 given in the specification.
CC The Polymorphisms are useful for studying the biological function of
CC ACHE as well as in identifying drugs targeting this protein for the
CC treatment of disorder related to its abnormal expression or function.
CC The polymorphic variants may also be used in screening for compounds
CC targeting ACHE to treat a specific condition or disease predicted to be
CC associated with ACHE activity e.g. neurological diseases (e.g.
CC Parkinson's disease and Alzheimer's disease), cancer, leukaemia,
CC and tumours. The ACHE gene maps to human chromosome 7q22. The
CC present sequence is the ACHE gene.

XX
SQ Sequence 12113 BP; 2215 A; 3836 C; 3766 G; 2296 T; 0 other;

Query Match 69.4%; Score 1198; DB 24; Length 12113;
Best Local Similarity 81.8%; Pred. No. 1.9e-243;
Matches 1554; Conservative 0; Mismatches 0; Indels 346; Gaps 1;

QY 1 atgaggccccgcagtgctgtgcacacgccttcctggtctcccaactcttctcctc 60
Db |||||
QY 4095 atgaggccccgcagtgctgtgcacacgccttcctggtctcccaactcttctcctc 4154
Db |||||
QY 61 ctctctggtcctcctgggtgaggagtggggcgctgagggccggaggatgcagagctgctg 120
Db |||||
QY 4155 ctctctggtcctcctgggtgaggagtggggcgctgagggccggaggatgcagagctgctg 4214
Db |||||
QY 121 gtgacggtgcgtggggccggctgcggggcattgcctgaagacccccggggccctgtc 180
Db |||||
QY 4215 gtgacggtgcgtggggccggctgcggggcattgcctgaagacccccggggccctgtc 4274
Db |||||
QY 181 tctgcttctcctgggcatcccccttttgcggagcaccacctgggacccccgcgtcttctgcc 240
Db |||||
QY 4275 tctgcttctcctgggcatcccccttttgcggagcaccacctgggacccccgcgtcttctgcc 4334
Db |||||
QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacacaccttccagagtgtctgc 300
Db |||||
QY 4335 ccggagcccaagcagccttggtcaggggtggtagacgctacacaccttccagagtgtctgc 4394
Db |||||
QY 301 taccaatatgtggacacccctataccccaggttttggaggccacccagatgtggaaccccaac 360
Db |||||
QY 4395 taccaatatgtggacacccctataccccaggttttggaggccacccagatgtggaaccccaac 4454
Db |||||
QY 361 cgtgagctgagcagggactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
Db |||||
QY 4455 cgtgagctgagcagggactgcctgtacctcaacgtgtggacaccatacccccgccctaca 4514
Db |||||
QY 421 tccccacccccctgctcctgctggtatctatgggggtggcttctacagtgggcctctcc 480
Db |||||
QY 4515 tccccacccccctgctcctgctggtatctatgggggtggcttctacagtgggcctctcc 4574
Db |||||
QY 481 ttggacgtgtacatggccgcttcttgggtacaggccgagaggactgtgtgtgtccatg 540
Db |||||
QY 4575 ttggacgtgtacatggccgcttcttgggtacaggccgagaggactgtgtgtgtccatg 4634
Db |||||

Db 1354 tacctgttgaccacgggacctcaacctcctctggccccctggatgggggtcccccat 1413

QY 1435 ggctacgagatcgagttcatctttgggatccccctggaccctctctcgaaactacacggca 1494

Db 1414 ggctacgagatcgagttgtcttggattgcgctcaacgacagcctgaactacacgccc 1473

QY 1495 gaggagaaaatcttcgccagcgactgatcgatactggggccaaactttgccgcacaggg 1554

Db 1474 caagagaaggagctgagccgcagaaatgatcgctactgggccaacttcgccggacaggg 1533

QY 1555 gatcccaatgagccccgagagacccccaaaggcccccaaatggcccccgtaacagcggggct 1614

Db 1534 aacccccacca---ccctgccgataagagcggggacctggccccactacacgcctccccag 1590

QY 1615 cagcagtagcttagtctggacctgcggccgctgagaggtgcggcggggtgcgcgccccag 1674

Db 1591 ccgcagtagtccagctcaacaccccagccgctggccaccccaacccagcctgcgggccccag 1650

QY 1675 gcctgcgccttctggaaacgcgttctctcccccaaatgctcagcgccac 1721

Db 1651 atctgcgccttctggaaacacttctctcccccaactgctgaacgccac 1697

RESULT 12

AAZ49470

ID AAZ49470 standard; cDNA; 2416 BP.

XX

AC AAZ49470;

XX

DT 04-APR-2000 (first entry)

XX

DE Human wild type Butyrylcholinesterase (BCHE) protein encoding cDNA.

XX

KW Butyrylcholinesterase; BCHE allele; neurological disease; treatment; therapy; allelic variant; BCHE-K; apoE4 allele; neurofibromatosis; non-AD neurological disease; Alzheimer's disease; Huntington's disease; depression; amyotrophic lateral sclerosis; multiple sclerosis; stroke; Parkinson's disease; multi-infarct dementia; human; ss.

XX

OS Homo sapiens.

XX

FH Key

FT CDS

FT 130..1938

FT /*tag= a

FT /product= "Butyrylcholinesterase"

XX

PN WO9966072-A2.

XX

PD 23-DEC-1999.

XX

PF 16-JUN-1999; 99WO-IB01298.

XX

PR 16-JUN-1998; 98US-0089406.

XX

PA (NOVA-) NOVA MOLECULAR INC.

XX

PI Sevigny P, Wiebusch H, Schappert K;

XX

DR WPI;- 2000-126550/11.

DR P-PSDB; AAY44573.

XX

PT Prediction of drug efficacy for treating neurological diseases like Alzheimer's disease, neurofibromatosis, Huntington's disease -

XX

PS Claim 4; Fig 1; 37pp; English.

XX

CC The present cDNA sequence encodes wild type human butyrylcholinesterase (BCHE) protein. Determining BCHE allele status of a patient helps in predicting risk for neurological diseases, efficacy of therapy and determining treatment protocol. Presence of BCHE allelic variant, BCHE-K and apoE4 allele indicate patient's risk for having a neurological disease. This method enables treating Alzheimer's disease, depression,

CC neurofibromatosis, Huntington's disease, amyotrophic lateral sclerosis, multiple sclerosis, stroke, Parkinson's disease, multi-infarct dementia and other non-AD neurological diseases.

XX

SQ Sequence 2416 BP; 774 A; 431 C; 478 G; 733 T; 0 other;

Query Match 22.0%; Score 380; DB 21; Length 2416;

Best Local Similarity 55.2%; Pred. No. 5.6e-71;

Matches 790; Conservative 0; Mismatches 625; Indels 15; Gaps 2;

QY 146 ggggcatcgcctgaagacccccggggccctgtctctgtcttctcctgggcatccccctttg 205

Db 254 gaggatgaacttgacagtttttgggtggcacgggtaacagcctttcttggaaattccctatg 313

QY 206 cggagccacccatgggacccccgcctgtcctttctgcccacgggagcccaagcagccttggtcag 265

Db 314 cacagccacctcttggtagacttcgattcctcaaaagccacagctcttgaccaagtgtctg 373

QY 266 ggggtgtagacgtctacaaccttccagagtgctgtctgtaccaaatatgtggacacctatacc 325

Db 374 atatttggaatgcccaaaatatgcaaatcttctgtctgcagaacatagatcaaaagttttc 433

QY 326 caggttttgagggccaccgagatgtggaaaccccccaaccgtgagctgagcgaggactgcctgt 385

Db 434 caggcttccatggatcagagatgtggaaaccccccaacactgacctcagtgaaagactgtttat 493

QY 386 acctcaacgtgtggacaccataccccccggcctacatccccccacccccctgtcctcgtctgga 445

Db 494 atctaaatgtatggattccagcacctaaacccaaaaaatgcc---ctgtatgatatagga 550

QY 446 tctatgggggtggtctctacagtggggcctcctcctcttgacgtgtacgattgcccgttct 505

Db 551 tttatggtggtgttttcaaacctggaacatcatctttacatgtttatgtatggtgcaagtctc 610

QY 506 tggtagcggccgagaggactgtgtggtgtccatgaactaccgggtggagcctttggct 565

Db 611 tggctcgggttgaaagagtattgttagtgtcaatgaactatagggtgggtgccctaggat 670

QY 566 tcctggccctgcggggagccgagagggccccgggccaatgtgggtcctcctggatcagagggc 625

Db 671 tcttagctttgccaggaaaatcctgaggtccagggaacatgggttatttgcatacaacagt 730

QY 626 tggccctgcagtggtgtaggagagaacgtggcagccttcggggtgacccgcacatcagtg 685

Db 731 tggctcttcagtggttcaaaaaatatagcagcctttggtggaaatccctaaaaagtgtaa 790

QY 686 cgctgtttggggagagcgcgggagccgctcctcgtgggcatgcacctgctgtccccgccca 745

Db 791 ctctctttggagaaagtgcaggagcagcttcagttagcctgcatcttctcctcctgga 850

QY 746 gccggggcctgtttccacagggccgtgctgcagagcggtgcccccaatggacctggggcca 805

Db 851 gccattcattgttccacagagccattctgcaaatggtatcctttaatgctccttgggcgg 910

QY 806 cgggtggcatgggagaggcccgctgcaggggccagcagctggcccaccttgtgggctgtc 865

Db 911 taacatctctttatgaagctagggaacagaaacgttgaaacttagctaatt----- 959

QY 866 ctccagggcgactggtgggaatgacacagagctggttagcctgccttcggacacgaccag 925

Db 960 -gactggttgcctctagagagaatgagactgaaataatcaagtgtcttagaaataaagatc 1018

QY 926 cgcaggtcctggtgaaccacgaatggcacgtgtgtgcctcaagaagcgtcttccggttct 985

Db 1019 cccaagaaattcttctgaatgaagcatttgtgtccccctatgggactcctttgtcagtaa 1078

QY 986 ccttcgtgcctgtggtagatggagacttctcctcctcagtgacacccccagagccctcatcaacg 1045

Db 1079 actttggtccgacctggtggtgattttctcactgacatgccagacatatatttgaaac 1138

QY 1046 cgggagacttccacggcctgcagggtgctggtggtggtggtgaaggatgagggtcgtatt 1105

Db 1139 ttggacaatttaaaaaaacccagattttggtgggtgtaataaagatgaagggacagctt 1198

QY 1106 ttctggtttacggggcccgagcttcagcaaaagacaacagagtctctcatcagccgggccc 1165

Db 1199 ttttagtctatggtcctcctggttcagcaaaagataacaatagatcataactagaaaaag 1258

QY 1166 agttcctggccgggtgcgggtcggtgtcccccaggtaaagtacctggcagccgagggctg 1225

Db 1259 aatttcaggaaggtttaaaaatatattttccaggagtgagtgagtttgaaaaggaatcca 1318

QY 1226 tggctcctgcattacacagactggctgcatacccgagagcccgccagcctgagggagggccc 1285

Db 1319 tcctttttcattacacagactgggtagatgatcagagacacctgaaaactaccgtgagggcct 1378

QY 1286 tgagcgtatggtgggcacacacaatgtcgtgtgcccgcgtggcccagctggctgggcgac 1345

Db 1379 tgggtgatgtgtgtgggattataatttcataatgcctgccttgaggttcaccaagaagt 1438

QY 1346 tggctgccaggggtcccggtctacgcctacgtcttttgaacacccgtgcttccagctct 1405

Db 1439 tctcagaatggggaaataatgccttttctactattttgaacacccgatcctccaaacttc 1498

QY 1406 cctggccctgtggtgggtgccccacggctacgagatcgagttcatctttgggatcc 1465

Db 1499 cgtggccagaatggatggagtgatgctgctggtgctatgaaattgaattgtctttggttac 1558

QY 1466 ccctggaccctctcgaactacacggcagagagagaaaattcttcgccagcgactgatgc 1525

Db 1559 ctctggaaaagaagagataattacacaaaagccgaggaataatttgatgataccatagtga 1618

QY 1526 gatactgggccaactttgccgcacaggggatcccaatgagccccgagac 1575

Db 1619 aacggtggcaaatttgcaaaatatgggaatcccaatgagactcagaac 1668

RESULT 13
AAZ49471

ID AAZ49471 standard; cDNA; 2416 BP.

XX

AC AAZ49471;

XX

DT 04-APR-2000 (first entry)

DE Human Butyrylcholinesterase-K (BCHE-K) protein encoding cDNA.

XX

KW Butyrylcholinesterase-K; BCHE-K; BCHE allele; neurological disease; therapy; treatment; allelic variant; apoE4 allele; neurofibromatosis; non-AD neurological disease; Alzheimer's disease; Huntington's disease; depression; amyotrophic lateral sclerosis; multiple sclerosis; stroke; Parkinson's disease; multi-infarct dementia; human; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 130..1938

FT /*tag= a

FT /product= "Butyrylcholinesterase-K"

FT replace(1828, G)

FT /*tag= b

XX WO9966072-A2.

PN

XX

PD 23-DEC-1999.

XX

PF 16-JUN-1999; 99WO-IB01298.

XX

PR 16-JUN-1998; 98US-0089406.

XX

PA (NOVA-) NOVA MOLECULAR INC.

XX

PI Sevligny P, Wiebusch H, Schappert K;

XX

XX WPI; 2000-126550/11.

DR

DR P-PSDB; AAY44574.

XX

PT Prediction of drug efficacy for treating neurological diseases like Alzheimer's disease, neurofibromatosis, Huntington's disease -

XX

PS Claim 13; Fig 2; 37pp; English.

XX

CC The present cDNA sequence encodes human polymorphic variant butyrylcholinesterase-K (BCHE-K) protein. BCHE-K is an allelic variant of BCHE. Determining BCHE allele status (homozygous or heterozygous) of a patient helps predicting risk of neurological diseases, efficacy of therapy and determining treatment protocol. BCHE-K and apoE4 allele status also indicate patient's risk for having a neurological disease. This method enables treating Alzheimer's disease, Huntington's disease, depression, neurofibromatosis, amyotrophic lateral sclerosis, stroke, multiple sclerosis, Parkinson's disease, multi-infarct dementia and other non-AD neurological diseases.

SQ Sequence 2416 BP; 775 A; 431 C; 477 G; 733 T; 0 other;

Query Match 22.0%; Score 380; DB 21; Length 2416;

Best Local Similarity 55.2%; Pred. No. 5.6e-71;

Matches 790; Conservative 0; Mismatches 625; Indels 15; Gaps 2;

QY 146 ggggcattcgccctgaagacccccggggccctgtctctgtcttctgggcaccccccttg 205

Db 254 gagggatgaacttgacagtttttgggtggcacggtaacagccttctcttgaattccctatg 313

QY 206 cggagccacccatgggaccccgctgtttctgcacccggagcccaagcagccttgggtcag 265

Db 314 cacagccacctcttggtagacttcgattcaaaaagccacagtcctctgaccaagtggctg 373

QY 266 ggggtggtagacgtacaaccttcacagagtgtctgtctacacaatatgtggacacccctatacc 325

Db 374 atatttgggaatgccacaaaatatgcaaatcttctgtcagaacatagatcaaaagttttc 433

QY 326 caggttttgagggcacccgagatgtggaaccccccaacctgagctgagcgaggactgcctgt 385

Db 434 caggcttccatggatcagagatgtggaaccccaaacacigacctcagtgaaagactgtttat 493

QY 386 acctcaacgtgtggacacaccataccccgggcctacatccccacccctgtcctcgtcttga 445

Db 494 atctaaatgtatggattccagcacctaaaccaaataatgccca--ctgtattgatataaga 550

QY 446 tctatgggggtggccttctacagtgggggccctcctcctcttgagcgtgtacgatggcgccttct 505

Db 551 tttatgggtggtgttttcaaaactggaaacatcatctttacatgttttatgatggcaagtttc 610

QY 506 tgggtacagggccgagagagactgtcgtggtgtccatgaactaccgggtgggagcctttggct 565

Db 611 tgggtcgggttgaaagagttattgtagtgtcaatgaactatagggtgggtgcccctaggat 670

QY 566 tcctggccctgccggggagccgagagagggccccgggcaatgtgggtctcctggatcagagggc 625

Db 671 tcttagcttggccaggaaatcctgaggtccagggaacatgggtttatttgatcaacagt 730

QY 626 tggccctgcagtggtgcagagaaacgtggcagccttcgggggtgacccgacatcagtga 685

Db 731 tggctcttcagtggttcaaaaaataatagcagccttgggtggaatacctaaaaagtgtaa 790

QY 686 cgctgtttgggagagcgcgggagccgcctcgtggtggcatgcacctgtctccccgccca 745

Db 791 ctctctttggagaaagtgcaggagcagcttcagttagccttgcatcttctctcctggaa 850

QY 746 gccggggcctgttccacagggcccgctgcgtgcagagcgggtgcccccaatggaccctggggcca 805

Db 851 gccattcattgttcaccagagccattctgcaaaagtggatcctttaatgctccttgggcgg 910

QY 806 cgggtgggcagtgggagagggcccgctgcgtgcagagcgggtgcccccaatggaccctggggc 865

Db 911 taacatctcttatgaagctaggaacagaacagcttgaaacttagctaaatt----- 959

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:25:56 ; Search time 6064.26 Seconds
(without alignments)
3839.256 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atgagcccccgcagtgctct.....aattgctcagcgccacctga 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_gss:*
13:	em_gss_hum:*
14:	em_gss_inv:*
15:	em_gss_pln:*
16:	em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1281	74.3	2080	11	BC001541 Homo sapi
2	764.2	44.3	990	10	BI489087
3	750.4	43.5	931	10	BI488594
4	698.6	40.5	780	10	BI753192
5	688.6	39.9	822	9	AI190022
6	638.8	37.0	752	10	BI917862
7	638.2	37.0	781	10	BG705869
8	623.8	36.2	786	10	BG703349
9	618.4	35.8	732	10	BG708331
10	615.4	35.7	738	10	BI667712
11	588.8	34.1	702	10	BI915019
12	560.2	32.5	589	9	AI654042
13	558	32.3	562	10	BE466391
14	555.2	32.2	689	10	BG395920
15	547	31.7	722	10	BE535653
16	538.8	31.2	555	9	AI769167
17	529.8	30.7	534	9	AI332425

C 18	521.6	30.2	577	9	AI802781	AI802781 wf18c08.x
C 19	494.6	28.7	549	9	AI143271	AI143271 qc04e07.x
C 20	484.6	28.1	574	9	AI207928	AI207928 ap14d04.x
C 21	479.6	27.8	533	9	AA406567	AA406567 zv11a11.s
C 22	476.4	27.6	478	9	AI082805	AI082805 ox74d09.x
C 23	474	27.5	507	9	AW196326	AW196326 xm31h05.x
C 24	469.8	27.2	474	9	AI761618	AI761618 wg66f01.x
C 25	467.8	27.1	483	9	AI761619	AI761619 wg66f02.x
C 26	461	26.7	483	9	AI761609	AI761609 wg66e02.x
C 27	454.4	26.3	456	9	AI081892	AI081892 ox77b06.x
C 28	454.4	26.3	932	10	BF235425	BF235425 602025226
C 29	444.4	25.8	625	10	BI464924	BI464924 603207402
C 30	432.4	25.1	485	9	AA447461	AA447461 zw89g10.r
C 31	432	25.0	514	9	AA032228	AA032228 zf01f07.s
C 32	427.8	24.8	489	9	AI091215	AI091215 oo22c11.x
C 33	424.4	24.6	586	9	AA446656	AA446656 zw89g10.s
C 34	420.6	24.4	523	10	BF191608	BF191608 239241 MA
C 35	417	24.2	525	10	BM256199	BM256199 518509 MA
C 36	414.6	24.0	506	10	BE721978	BE721978 189843 MA
C 37	412.6	23.9	550	10	BE808194	BE808194 213342 MA
C 38	412.6	23.9	558	10	BF074683	BF074683 222160 MA
C 39	402.4	23.3	405	9	AW050645	AW050645 wz19f11.x
C 40	401.6	23.3	495	9	AA411703	AA411703 zv11a11.r
C 41	401	23.2	821	10	BE540575	BE540575 601067212
C 42	397.2	23.0	662	9	BB629080	BB629080 BB629080
C 43	396.4	23.0	556	9	AI332839	AI332839 qp96g08.x
C 44	382	22.1	504	10	BM286130	BM286130 526062 MA
C 45	377.2	21.9	396	10	H21132	H21132 yn65b07.r1

ALIGNMENTS

RESULT 1

BC001541

LOCUS BC001541 2080 bp mRNA linear HTC 31-JAN-2002

DEFINITION Homo sapiens, clone IMAGE:3453362, mRNA.

ACCESSION BC001541

VERSION BC001541.1 GI:14705895

KEYWORDS HTC.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2080)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (21-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 4 Row: e Column: 19

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7710156

This clone has the following problem: no polyA-tail.

Location/Qualifiers

1. .2080

/organism="Homo sapiens"

FEATURES

source

```

/db_xref="taxon:9606"
/clone="IMAGE:3453362"
/tissue_type="Placenta,
/clone_lib="NIH_MGC_10"
/lab_host="DH10B"

```

```

BASE COUNT      319 a      692 c      651 q      418 t
/note="Vector: pCMV-SPORT6"

```

Query Match	74.3%;	Score 1281;	DB 11;	Length 2080;
Best Local Similarity	82.6%;	Pred. NO. 1e-234;		
Matches 1637;	Conservative	0;	Mismatches	0;
			Indels	346;
			Gaps	1;

QY	1	atgagggccccgcagtgctgtgcacacgccttccctggcttcccactcttctcctc	60
Db	98	ATGAGGCCCCCGCAGTGTCTGTGCACACGCCCTTCCCTGGCTTCCCCACTCTCTCCTC	157
QY	61	ctcctctggctcctgggtggagagtgggggctgagggccgggaggatgcagagctgctg	120
Db	158	CTCCTCTGGCTCCTGGGTGGAGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG	217
QY	121	gtgacggtgctggtggggccggtcgcggggcattcgccctgaagacccccccggggccctgctc	180
Db	218	GTGACGGTGCCTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCCCGGGGGCCCTGTC	277
QY	181	tctgctttcctgggcatcccccttttgcggggcaccccatgggaccccgctcgctttctgcca	240
Db	278	TCTGCTTTCTTGGGCATCCCCCTTGGCGGAGCCACCCATGGGACCCCGTCGCTTCTTGCCA	337
QY	241	ccggagcccaagcagcccttgctcaggggtggtagacgctacaaccttcacagagtgtctgc	300
Db	338	CCGGAGCCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC	397
QY	301	taccaatatgtggacacccctatccccaggttttgagggcacccgagatgtggaacccccaac	360
Db	398	TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATCTGGAACCCCAAC	457
QY	361	cgtgagctgagcgaggactgcctgtacctcaactcaacgtgtggacaccatacccccgccctaca	420
Db	458	CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCGGCCCTACA	517
QY	421	tccccacccctgtcctcgtctggatctatgggggtggtctctacagtggggcctcctcc	480
Db	518	TCCCCACCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCTCTCC	577
QY	481	ttggacgtgtacgatggccgctcttgggtacctcaacgtgtgacagggagactgtgctgggtgccatg	540
Db	578	TTGGACGTGTACGATGGCCCGCTTCTTGGTACAGGCCGAGAGGACTGTGTGCTGGTGTCCATG	637
QY	541	aactaccgggtggagcctttggcttccctgcctgcccctgcccgggagccgagagccccgggc	600
Db	638	AACTACCGGTGGGAGCCCTTGGCTTCTTGGCCCTGCCCGGGAGCCGAGAGGCCCCCGGGC	697
QY	601	aatgtgggtctcctggatcagaggctggccctgcagtggtgagagagaacagtggcagcc	660
Db	698	AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC	757
QY	661	ttcgggggtgaccccgacatcagtgacgctgtttggggagagcgccgggagccgctcggtg	720
Db	758	TTCCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGAGCCCGCTCGGTG	817
QY	721	ggcatgcacctgctgtcccccgccccagccggggcctgtttccacagggccgtgctgcagagc	780
Db	818	GGCATGCACCTGCTGTCCCCCGCCAGCCCGGGCCCTGTTCCACAGGGCCGTGCTGCAGAGC	877
QY	781	ggtgcccccaatggaacctgggccacggtgggcatgggagagggcccgctgcagggccacg	840
Db	878	GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCGCTGCAGGGCCACG	937
QY	841	cagctggcccaacctgtgggctgtcctccaggcgccactggtgggaatgacacagagctg	900
Db	938	CAGCTGGCCCCACCTTGTGGGCTGTCTCTCCAGCGGCACCTGTTGGGAATGACACAGAGCTG	997

Plate: LLAM11479 row: 1 column: 06
High quality sequence stop: 771.
Location/Qualifiers
1. .931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5191829"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

BASE COUNT 152 a 294 c 306 g 179 t
ORIGIN

Query Match 43.5%; Score 750.4; DB 10; Length 931;
Best Local Similarity 96.9%; Pred. No. 1.7e-133;
Matches 882; Conservative 0; Mismatches 16; Indels 12; Gaps 11;

QY 130 cgtggggcggtgcgtgcgggcatcgcctgaagacccccgggggcccctgtctctgtcttc 189
Db 1 CGTGGGGCGCGCTGCGGGGCATTCGCCTGAAGACCCCGGGGGCCCTGTCTCTGCTTTC 60

QY 190 ctgggcatcccccttgcggagccaccccatgggacccccgcgtcttctgccaccggagccc 249
Db 61 CTGGGCATCCCTTTGCGGAGCACCCCATGGGACCCCGCTGCTTCTGCCACCGGAGCCC 120

QY 250 aagcagccttggtcaggggtgtagacgctacaaccttccagagtgtctgctaccaatat 309
Db 121 AAGCAGCCTTGTCAGGGGTGTAGACGCTACAACCTTCCAGAGTGTCTGTACCAATAT 180

QY 310 gtgacacccctatacccaggttttaggggcaccgagatgtggaaccccccaaccgtgagctg 369
Db 181 GTGCACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAACCGTGAGCTG 240

QY 370 agcgaggactgcctgtacctaacagtgtggacaccataccccccggcctacatcccccaacc 429
Db 241 AGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATAACCCCGGCCTACATCCCCCACC 300

QY 430 cctgtcctcgtctggtatctatgggggtggtcttctacagtgggggcctcctcttg-gacgt 488
Db 301 CCTGTCTCTGCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCCTCCTCTTGTGACGT 360

QY 489 gtacgatggccgcttctgttacagggccgagaggaactgtgctggtgccatgaactacgg 548
Db 361 GTACGATGGCCGCTTCTTGTGTACAGGCCGAGAGGACTGTGCTGGTGTCCATGAACATCCG 420

QY 549 ggtggagccttggcttccttgccctgcgcctgcggggagccgagagccccgggcaatgtggg 608
Db 421 GGTGGAGCCTTTGGCTTCTTGCTTCCCGGGGAGCCGAGAGGCCCGGGCAATGTGGG 480

QY 609 tctcctggatcagaggctggccctgcagtggggtgcaggagaacgtggcagccttcgggggg 668
Db 481 TCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCCTTCGGGGGG 540

QY 669 tgacccgacatcagtgacgctgttttggggagagcgcggagccgcctcgtggtgggcatgca 728
Db 541 TGACCCGACATCAGTGACGCTGTTTGGGGAGAGCCGGAGCCGCCCTCGTGGGCATGCA 600

QY 729 cctgctgtccccccagcccgggg-cctgttccacagggccgctgctgcagagc-ggtgcc 786
Db 601 CTTGCTGTCCCCCGCCAGCCGGGGCCCTGTTCCACAGGGCCGTGCTGCAGAGCGGGTCC 660

QY 787 cccaatggaccctgggccacggtgggcatgggagaggcccgctcgagggccac--geagc 844
Db 661 CCCAATGGACCCCTGGGCCACCGTGGGCATGGGAGAGGGCCCGTCGCAGGGGCCACAGCAGCT 720

QY 845 tggccacacctgtggtggtgt-cttcacagcggcactggtgggaatgacacagagct-ggt 902
Db 721 TGGCCACACCTTGTGGGCTGTCCCTCCAGGGGCACCTGTTGGGAATGACACAGAGCTGGT 780

QY 903 agcctgcc-ttcggacacgacc-agcgaggtcctcctggtgaaccacgaatggcacgtgctg 960
Db 781 AGCCTGCCTTTCGGACACGACCAAGCGCAGGTCTCTGTTGAACACAGAAATGGCACGTCGTG 840

QY 961 cctcaagaagcgtcttcctcggttctccttcctcgtgcctgtgtagatggagacttccctcagt 1020
Db 841 -CTCAAGAAAGCGTCTTCCGG-TCTCCTTCGTGCCTGT-GTCCATTGCGACTTCCCTCAGG 897

QY 1021 gacacccccag 1030
Db 898 AAACCCACAG 907

RESULT 4
BI753192
LOCUS 603026066F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196256 5',
DEFINITION mRNA sequence.
ACCESSION BI753192
VERSION BI753192.1 GI:15744770
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 780)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11491 row: d column: 17
High quality sequence stop: 780.

FEATURES
source Location/Qualifiers
1. .780
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5196256"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

BASE COUNT 111 a 255 c 260 g 154 t
ORIGIN

Query Match 40.5%; Score 698.6; DB 10; Length 780;
Best Local Similarity 98.4%; Pred. No. 1.3e-123;
Matches 717; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

QY 1 atgaggcccccgagtgctgtcgtgcacagccttcctcgttcctccactcctctc 60
Db 48 ATGAGGCCCCCGCAGTGTCTGTGTGCACACGCTTTCCTGGCTTCCCCACTCCTCTCCTC 107

QY	61	ctcctcgtcctcgtggtgagagtggtgggctgagggcggtgagggccgggaggatgcagagctgctg	120
Db	108	CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGCTGAGGCGCGGAGGATGCAGAGCTGCTG	167
QY	121	gtgacggtggtggtggccggtcggtcggtggtgattcgctgaagaccccccggtggccctgtc	180
Db	168	GTGACGGTGCCTGGGGGCGCGGTGCGGGGCATTGCGCTGAAGACCCCCCGGGGCCCTGTC	227
QY	181	tctgctttctggtggtcctcccttttgcggagccaccctatgggacccctgcgtctttctgcc	240
Db	228	TCTGCTTTCCTGGGCATCCCCCTTTCGGGAGCACCCATGGGACCCCGTCGCTTCTGCCA	287
QY	241	ccggagcccaagcagcccttggtcagggttggtagacgtacaaaccttcacagagtgtctgc	300
Db	288	CCGGAGCCCAAGCAGCCTTGGTCAGGGTGGTAGACGCTACAAACCTTCCAGAGTGTCTGC	347
QY	301	taccaatatgtggacaccctataccccaggttttgaggccagagatgtggaaccccaac	360
Db	348	TACCAATATGTGGACACCCTATACCCAGGTTTGAGGGCACCGAGATGTGAACCCCAAC	407
QY	361	cgtgagctgagcagagactgcctgtacctcaacgtgtggacaccataccccggcctaca	420
Db	408	CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTTGGACACCATACCCCGGCTACA	467
QY	421	tccccaccctctcctcgtctggtatctatgggggtggcttctacagtggggcctcctcc	480
Db	468	TCCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC	527
QY	481	ttggacgtgtacgatggcgcgtcttctgtgtacagggccgagagagactgtgctggtgtccatg	540
Db	528	TTGGACGTGTACGATGGCCGCTTCTTGGTGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG	587
QY	541	aactaccggtggtgagcccttggcttcctggtccctgcctgcgggagccgagagccccgggc	600
Db	588	AACTACCGGTTGGAGCCCTTGGCTTCTTGGCTACAGGCCGAGAGGACTGTGCTGGTGTCCATG	647
QY	601	aatgtgggtcctcgtgatcagaggtggcctgcagtggtgcaggagaacgtggcagcc	660
Db	648	AATGTGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC	707
QY	661	ttcgggggtgaccgcgacatcagtgacgctgttttggggagagcgcgggagccgcctcggtg	720
Db	708	TTCGGGGGTGACCCGACATCAGTGACGCTGT--GGGGAGAGCGCGGAGCCGCTCGGTG	764
QY	721	ggcatgcac	729
Db	765	GGCATGCCC	773
RESULT 5			
AI190022/c			
LOCUS			
DEFINITION			
AI190022 822 bp mRNA linear EST 28-OCT-1998			
qd35el2.x1 Soares_fetal_heart_NbHH19W Homo sapiens CDNA clone			
IMAGE:1731502 3' similar to gb:M55040 ACETYLC HOLINESTERASE			
PRECURSOR (HUMAN);, mRNA sequence.			
AI190022			
AI190022.1 GI:3741231			
EST.			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
1 (bases 1 to 822)			
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
AUTHORS			
TITLE			
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
Tumor Gene Index			
JOURNAL			
Unpublished (1997)			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgaps-r@mail.nih.gov			
This clone is available royalty-free through LLNL ; contact the			
IMAGE Consortium (info@image.llnl.gov) for further information.			
Insert Length: 1019 Std Error: 0.00			
Seq primer: -40ml3 fwd. ET from Amersham			

FEATURES		High quality sequence stop: 469.	
source		Location/Qualifiers	
		1..822	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:1731502"	
		/clone_lib="Soares_fetal_heart_NbHH19W"	
		/sex="unknown"	
		/dev_stage="19 weeks"	
		/lab_host="DH10B (ampicillin resistant)"	
		/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a	
		modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st	
		strand cDNA was primed with a Not I - oligo(dT) primer [5'	
		TGTACCAATCTGAAGTGGGACGCGCCGTCATCTTTTCTTTTCTTTT 3'],	
		double-stranded cDNA was size selected, ligated to Eco RI	
		adapters (Pharmacia), digested with Not I and cloned into	
		the Not I and Eco RI sites of a modified pT7T3 vector	
		(Pharmacia). Library went through one round of	
		normalization to a Cot = 5. Library constructed by	
		M.Fatima Bonaldo. This library was constructed from the	
		same fetus as the fetal lung library, Soares fetal lung	
		NBHL19W."	
BASE COUNT		151 a 271 c 245 g 150 t	5 others
ORIGIN			
		Query Match 39.9%; Score 688.6; DB 9; Length 822;	
		Best Local Similarity 93.8%; Pred. No. 1.1e-121;	
		Matches 769; Conservative 0; Mismatches 44; Indels 7; Gaps 5;	
QY	833	gggccacgcagctggccccacctgtgtggctgtcctccaggcgccactgg--tgggaaatga	890
Db	820	GGCCAAACGAAGTGGCCCCACTTGGTGGGTGTCTCCAGCGGGCATGGGTGGGGAATGA	761
QY	891	cacagagctggtagcctgccttcggacacgaccagcgca-ggtcctggtgaaccacgaat	949
Db	760	CACAAAAGTGGTAAGCTGCCTTCGGACACGACACGAGGCGAGGCTCTTGGTGAACAGAA	701
QY	950	ggcacgtgctgcctcaag-aaagcgtcttcctcggtctctcc--ttcgtgcctgtggttagatg	1006
Db	700	GGCACGTGCTGCTCNAAGAAAACGCTCTTCCGGGTCTCCCTTCNGTGCCTGTGGTAGATG	641
QY	1007	gagacttctcagtgacacccccagagg-ccctcatcaacgcgggagaccttcacggcctg	1065
Db	640	NAGACTCCTCAAGTGACACGCCAGAGGCCCTCATCAACGCGGAGACTTCACGGGCGCTG	581
QY	1066	caggtgctggtggtggtgaaggatgagggctcgtatttctggtttacggggcccca	1125
Db	580	CAGGTGCTGGTGGTGGTGAAGGATGAGGGCTCGGTATNTTCTGGTTTACGGGGCCCCA	521
QY	1126	ggcttcagcaaacagacagtgctctcatcagcggggccgagttccttgccggggtgcgg	1185
Db	520	GGCTTCAGCAAAAGACAACAGAGTCTCTCATCAGCCGGGCGGAGTTTCTTGGCCGGGTGCGG	461
QY	1186	gtcgggggttccccaggtaagtgcctggcagccagaggctgtggtcctgcattacacagac	1245
Db	460	GTCGGGGTTCCCCAGGTAAGTGACCTGGCAGCCGAGGCTGTGGTCTGCATTACACAGAC	401
QY	1246	tggctgcataccccgaggaccccgccacgcctgaggaggccctgagcgaatgtgtggtgcgac	1305
Db	400	TGGCTGCATCCCCGAGGACCCGGCACGCCTGAGGAGGCCCTGAGCGATGTGTGGGCGGAC	341
QY	1306	cacaatgtcgtgtgccccgtggccccagctggctggggcagctggctgccccaggtgcctccg	1365
Db	340	CACAATGTGCTGTGCNCCGTGGCCCCAGCTGGCTGGCGGACTGGTGTGCCAGGTTGCCCGG	281
QY	1366	gtctacgcctacgtcttttgaacaccgtgtcttccacgctctcctggccccctgtggtgggg	1425
Db	280	GTCTACGCCTACGCTCTTGAACACCCGTGCTTCCACGCTCTCCTGGCCCCCTGTGGATGGG	221
QY	1426	gtgccccacggctacgagatcgagttcatcttttgggatccccctggacccccctctcgaaac	1485
Db	220	GTGCTCACGGCTACGAGATCGAGTTCATCTCTTTGGGATCCCCCTTGAGCCCCCTCTCGA	161


```
QY 1486 tacacggcagaggagaataatcttcgccccagcgactgatcgatactgggccaactttgcc 1545
|||||
Db 160 TACACGGCAGAGGAGAGAAATCTTCGCCACGCGACTGATCGATACTGGGCCAACTTTGCC 101

QY 1546 cgcacaggggatcccaatgagccccgagagaccccccaagaatggcccccgctacacg 1605
|||||
Db 100 CGCACAGGGGATCCCAATGAGCCCCGAGACCCCAAGSCCCCAACAATGGCCCCCGTACACG 41

QY 1606 gcgggggctcagcagtagcttagtctggacctgcggccgc 1645
|||||
Db 40 CGGGGGCTCAGCAGTACTTAGTCTGGACCTCGGSCGC 1

RESULT 6
BI917862 752 bp mRNA linear EST 17-OCT-2001
LOCUS 603183996F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248006 5',
DEFINITION mRNA sequence.
ACCESSION BI917862
VERSION BI917862.1 GI:16199790
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 752)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1625 row: p column: 23
High quality sequence stop: 752.
Location/Qualifiers
1..752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5248006"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
BASE COUNT 106 a 248 c 251 g 147 t
ORIGIN

Query Match 37.0%; Score 638.8; DB 10; Length 752;
Best Local Similarity 98.7%; Pred. No. 3.4e-112;
Matches 665; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 1 atgaggccccgcagtgctgtgcacacgccttcctcgcttccccactctctctc 60
|||||
Db 80 ATGAGCCCCCGCAGTGTGTGTGCACACGCCTTCCTGGCTTCCCCACTCCTTCTCCTC 139

QY 61 ctcctctggctcctggtaggagtgaggggctgagggccggaggatgcagagctgtg 120
|||||
Db 140 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 199
```

```
QY 121 gtgacgggtgcgtggggcgcgctgcggggcattcgctgaagacccccggggccctgtc 180
|||||
Db 200 GTGACGGTGC GTGGGGCGCGCTGCGGGGCATTGCGCTGAAGACCCCCCGGGGCGCTGTC 259

QY 181 tctgcttctcgtggcatccctcttgcggagccaccatgggaccccgctgcttctgtcca 240
|||||
Db 260 TCTGCTTTCTTCTGGGCATCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTTCTGCCA 319

QY 241 ccggagcccaagcagccttggtcagggggtgtagacgtacaaccttccagagtgctctgc 300
|||||
Db 320 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGTACAACCTTCCAGAGTGTCTGTC 379

QY 301 taccaatatgtggacacccctataccacagggttttgaggggcaccgagatgtggaaccccaac 360
|||||
Db 380 TACCAATATGTGGACACCTATACCCAGGTTTGTAGGGCACCGGAGATGTGGAAACCCCAAC 439

QY 361 cgtgagctgagcagaggactgcctgtacctcaacgtgtggacacccatcccccgctaca 420
|||||
Db 440 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCGCTACA 499

QY 421 tccccacccctgtcctgtctgtgatctatgggggtggtcttctacagtggggctcctcc 480
|||||
Db 500 TCCCCACCCCTGTCTCTGCTGTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCTCTCC 559

QY 481 ttggacgtgtacgatggcgcttcttgggtacaggccgagagactgtgctggtgtccatg 540
|||||
Db 560 TTGGACGTGTACGATGSCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 619

QY 541 aactaccgggtggagcctttggcttcctcctggccctgcgggggagcagagagggcccgggc 600
|||||
Db 620 AACTACCGGTGGGAGCCTTTGGCTTCTCTGGCCCTGCCGGGAGCCGAGAGGCCCGGGGC 679

QY 601 aa-tgtgggtctcctgatcagagaggtggccctgcagtggtgagtgaggaacgtggcagc 659
|||||
Db 680 AATTGTGGTCTCCTGGATCAGAGGCTGG-CCTGCACTGGGTGCAGGAGAACCGTGGCAG 738

QY 660 cttcgggggtgacc 673
|||||
Db 739 CTTGGGGGTGACC 752

RESULT 7
BG705869
LOCUS BG705869 781 bp mRNA linear EST 07-MAY-2001
DEFINITION 602669361F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792046 5',
mRNA sequence.
ACCESSION BG705869
VERSION BG705869.1 GI:13980645
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 781)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10669 row: j column: 15
High quality sequence stop: 781.
Location/Qualifiers
1..781
/organism="Homo sapiens"
```



```

/db_xref="taxon:9606"
/clone="IMAGE:4792046"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTT-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
110 a 260 c 262 g 149 t

```

Query Match 37.0%; Score 638.2; DB 10; Length 781;
Best Local Similarity 99.3%; Pred. NO. 4.4e-112;
Matches 662; Conservative 0; Mismatches 3; Indels 2;
Gaps 2;

Qy 1 atgaggccccgcagtgctgctgcacacgcctccctggcttccccactccttctctc 60

Db 113 ATGAGGCCCCGCAGTGTCTGTGCACACGCCCTTCCCTGGCTTCCCCACTCCTTCTCCTC 172

Qy. 61 ctccctctggctcctgggtggaggagtgggggctgagggccgggaggatgcagagctgctg 120

Db 173 CTCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 232

Qy 121 qtqacqqtqcqtqgqgqccqgcttcgacctgaagacccccqgqgccccctatc 180

Db 233 GTGACGGTGCGTGGGGCCGGCTGCGGGGCAATTCGCCCTGAAGACCCCGGGGGCCCTGTC 292

Qy 181 tctgcttcttgggcatccccctttgaggagccaccatgggagccccgtcgctttctgccca 240

D_b 293 TCTGCTTTCCTGGGCATCCCCCTTGGGGAGCCACCCCATGGGACCCCGTCGCTTCTGCCA 352

Qy 241 ccggagcccaagcagcccttggtcagggtggttagacgctacaaccttcagagtgctgc 300

Db 353 CCGGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGGTACAACCTTCCAGAGTGTCTGC 412

QY 301 taccaatatgtggacacccctataccccaggttttgagggcaccgagatgttggaacccaac 360

Db 413 TACCAATATGTGGACACCCCTATACCCAGGGTTTGAGGGCACCAGCATCTGCAACCCCAAC 413

0v 361 catgagctgaaggaagactgacctgtacctcaagacctgtgagacacacacatgccccccgacctgacg 420

Db 473 CGTGAGCTGAGCGGGACTGGCTGTACCTCACCCTGTGGACACCATATCCCCGGCTACA 532

QY 421 tccccccacctgtcctcgtctcggatctatggggggtggcctctctacagtgggggcctcctcc 480

Db 533 TCCCCA-CCTGTCCTCCTGCACTATGCCCCCTCCCTTATCACCTCCCCCTCCCC 501

0v 481 t t a g a c g t a t a c g a t a g c c g c t t c t t a g t a c a g c c g a g a g a c t a t a c t a g t a c c a t c 510

D_b 592 TTGGACGTGTACGATGGCCGGTTCTTGCTACAGGCCGAGAGGACTCTCCTCTCCATC 651

[illegible][illegible]

Qy 601 aatgtgggtctcctggatcagagggctggccctgcagtgggtgcaggagaaactggcagcc 660

[illegible][illegible]

771 771 771 771 771 771 771 771 771 771

RESULT 8
BG703349

LOCUS BG703349 786 bp mRNA linear EST 07-MAY-2001
 DEFINITION 602685108F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:4817881 5',
 mRNA sequence.
 ACCESSION BG703349
 VERSION BG703349.1 GI:13975591
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 786)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10718 row: 0 column: 02
 High quality sequence stop: 731.

Query Match 36.2%; Score 623.8; DB 10; Length 786;
Best Local Similarity 97.2%; Pred. No. 2.5e-109;
Matches 656; Conservative 0; Mismatches 17; Indels 2;
Gaps 2;

Qy 1 atgaggccccgcagtgtctgtgcacagccttccctggctccccactcctctctc 60

Db 113 ATGAGGCCCGCAGTGTCTGCTGCACACGCCCTTCCCCTGGCTTCCCCACTCCTTCTCTC

Qy 61 ctcctctggtggtgaggagtgagggtgagggccggaggatgcagagctgctg 120

db 173 CTCTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGGCTGCTG 232

Qy 121 gtgacggtcgctggggccggctgggggcattcgccctgaagacccccgggggcctgtc 180

Db 233 GTGACGGTGGCTGGGGGGCCGGCTGGGGGGCATTCGGCTGAAGACCCCGGGGGCCCTGTC 292

Qy 181 tctgctttcctgggcataccccctttgctggagccaccccatgggagccccgtgcgtttctgcca 240

db 293 TCTGCGTTTCCTGCGCATCCGCCCTTCCGAGCCACCCATCSCCACTCCCTTCCTCCCA 353

QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttcagagtgctgc 300

353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1

Qy 301 taccaatatgtggacacccctataccagggttttgagggcaccgagatgtggaaccccaac 360

Db	413	TACCAATATGTGGACACACCTATACCCAGGTTTTTGGGGCACCAGAGATGTGGAACCCCAAC	472
QY	361	cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccggcctaca	420
Db	473	CCTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGGCTACA	532
QY	421	tccccaccacctgtcctcgctggtgatctatgggggtggtcttctacagtgggacctctcc	480
Db	533	TCCCCCATCCCTGTCCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC	592
QY	481	ttggacgtgtacgatggcgcttcttctgtgtacagggccgagaggactgtgctggtgtccatg	540
Db	593	TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG	652
QY	541	aactaccgggtggagcctttggc-ttcttgccctgcggggagccgagagggccccggg	599
Db	653	AACTACCGGTGGGAGCCTTTGGCTTTCCTGGCCCTGCCGGGAGCCGAGAGG-CCCGGG	711
QY	600	caatgtgggtcctcctggatcagagggctggccctgcagtggtgcaggagaaacgtggcagc	659
Db	712	CAATGTGGGTCTCCTGGATCACACGCTGGCCTGCAGGTGGGTGCCGGAGAACCGTGGCAG	771
QY	660	cttcggggggtgaccc	674
Db	772	CTTCGGGGGTGACCC	786
RESULT	9		
BG708331			
LOCUS			
DEFINITION	602672268F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:4795048 5',	732 bp	mRNA linear EST 07-MAY-2001
ACCESSION	BG708331		
VERSION	BG708331.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 732)		
JOURNAL	NIH-MGC http://mgi.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10677 row: g column: 17 High quality sequence stop: 732.		
FEATURES	Location/Qualifiers		
source	1. .732		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4795048"		
	/clone_lib="NIH_MGC_96"		
	/tissue_type="hypothalamus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	101 a 250 c 239 g 142 t		

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11791 row: f column: 18
High quality sequence stop: 736.

FEATURES
source
1. .738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312273"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ; brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: Sali-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 107 a 231 c 248 g 152 t
ORIGIN

Query Match 35.7%; Score 615.4; DB 10; Length 738;
Best Local Similarity 97.6%; Pred. No. 9.9e-108;
Matches 657; Conservative 0; Mismatches 11; Indels 5; Gaps 3;

Qy 1 atgagccccgcagtgctgtcgtcacacgccttccctggcttccccactcttctctc 60
|||||
Db 62 ATGAGGCCCCGAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCACTCTTCTCCTC 121

Qy 61 ctccctgtgctcctgggtgagagtggtgggctgagggccggaggatgcagagctgctg 120
|||
Db 122 CTCGTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 181

Qy 121 gtgacggtgcgtggggccggctggggggcattcgccctgaagacccccgggccccctgtc 180
|||||
Db 182 GTGACGCTGCGTGGGGCCGGCTGCGGGGCATTTCGCCTGAAGACCCCCGGGGCCCTGTC 241

Qy 181 tctgctttcctgggcatccccctttgcggagccaccatgggacccccgtcgctttctgcc 240
|||||
Db 242 TCTGCTTCTCTGGGCATCCCCCTTTGCGGAGCCACCATTGGGACCCCGTCGCTTCTGCCA 301

Qy 241 ccggagcccaagcagccttggtgcaggggtggtagacgctacaacctccacagagtgtctgc 300
|||||
Db 302 CCGGAGCCCAAGCAGCCTTGGTTCAGGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 361

Qy 301 taccatatgtggacaccctataccagggttttgagggcaccagagatgtggaaccccaac 360
|||||
Db 362 TACCAATATGTGGACACCCTATACCCAGGTTTGTAGGGCACCGAGATGTGAACCCCAAC 421

Qy 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
|||||
Db 422 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATACCCCGGCTACA 481

Qy 421 tccccaccctgtcctcgtctggtatgtgggtggtggtcttctacagtggggcctcctcc 480
|||||
Db 482 TCCCCCACCCTGTTCATCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCCTCCTCC 541

Qy 481 ttggacgtgtacgatggccgcttcttggtacagggccgagagagactgtgc-tggtgtccat 539
|||||
Db 542 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTTGGTGTCCAT 601

Qy 540 gaactacgggtgggagcc---tttggttctctggccctgcgggggagcccgaga-ggccc 595
|||||
Db 602 GAACTACGGGTGGGAGAGCCTTTTGGCTTCTTGGCCCTGCTGGGGCCCTGCCGGGAGCCGAGAGGGCCC 661

Qy 596 cgggcaatgtgggtctctcctggatcagaggctggccctgcagtggtgaggaacgtgg 655

Db 662 CGGGCAATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAACTGG 721
|||||
Qy 656 cagccttcggggg 668
|||||
Db 722 CAGCCTTCGGGGG 734

RESULT 11
BI915019 702 bp mRNA linear EST 16-OCT-2001
LOCUS 603177204F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5241624 5',
DEFINITION mRNA sequence.
ACCESSION BI915019
VERSION BI915019.1 GI:16179103
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 702)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM11609 row: g column: 01
High quality sequence stop: 702.

FEATURES
source
1. .702
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5241624"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ; brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

BASE COUNT 99 a 240 c 226 g 137 t
ORIGIN

Query Match 34.1%; Score 588.8; DB 10; Length 702;
Best Local Similarity 99.4%; Pred. No. 1.2e-102;
Matches 612; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 atgagccccgcagtgctgtcgtgcacacgccttccctggcttccccactctctctc 60
|||||
Db 87 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCACTCTTCTCCTC 146

Qy 61 ctccctgtgctcctgggtggaggagtggtgggctgcctgaagacccccggggccctgtc 120
|||||
Db 147 CTCCTCTGGCTCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 206

Qy 121 gtgacggtgcgtggggccggctgctggggcattcgccctgaagacccccggggccctgtc 180
|||||
Db 207 GTGACGCTGCTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGCCCTGTC 266

Qy 181 tctgctttcctggggcatccccctttgctgggagccacccatgggacccccgtcttctgcc 240

||||| TCTGCTTCTGGGATCCCTTTGGGGAGCACCCATGGGACCCCGTCGCTTCTGCCA 326

QY 241 ccggagcccaagcgccttggtcaggggtggtagacgctacaaccttcagagtgctgc 300

Db 327 CCGAGCCCAAGCAGCCTTGGTCAGGGGTGTAGACGCTACACCTTCAGAGTGTCTGC 386

QY 301 taccaatatgtgacacccctataccagggttttgagggcaccgagatgtggaacccaac 360

Db 387 TACCAATATGTGCACACCTATACCCAGGTTTGAGGGCACCGAGATGTGGAACCCCAAC 446

QY 361 cgtgagctgagcaggactgcctgtacctcaacgctgtggacaccatacccccgccctaca 420

Db 447 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGCTGTGGACACCATACCCCGGACTACA 506

QY 421 tccccacccctgctcctcgtcgtggtatctatgggggtggtctctacagtgggacctctcc 480

Db 507 TCCCCAACCCCTGCTCCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 566

QY 481 ttggacgtgtacgatggccgcttcttgggtacaggcccgag-aggaactgtcgtggtgccat 539

Db 567 TTGGACGTGTACATGGCCCGCTTCTTGTGTACAGGCCGAGCAGGACTGTGCTGGTGTCCAT 626

QY 540 gaactaccgggtgggagcccttggcttccctggccctgccggggagccgagagggcccc-gg 598

Db 627 GAACTACCGGGTGGGAGCCTTTGGCTTCTGTGGCCTTCCGGCCCTGCCGGGGAGCCCGGGG 686

QY 599 gcaatgtgggtcctcct 614

Db 687 GCAATGTGGGTCTCCT 702

RESULT 12

AI654042/c

LOCUS

DEFINITION

ty61a12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2283550 3' similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);, mRNA sequence.

AI654042

AI654042.1 GI:4738021

EST.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 589)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/bbrp/image/image.html

Insert Length: 850 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 461.

Location/Qualifiers

1..589

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2283550"

/clone_lib="NCI_CGAP_Kid11"

/lab_host="DH10B"

/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was

FEATURES

source

prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 106 a 193 c 186 g 101 t 3 others

ORIGIN

Query Match 32.5%; Score 560.2; DB 9; Length 589;

Best Local Similarity 99.3%; Pred.No. 3.2e-97;

Matches 562; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1055 tcacagccctgcaggtgctggtggtggtgaaggatgagggctcgtatttctggttt 1114

Db 566 TCACCGGCCTGCAGGTGCTGGTGGGTGTGTGAAGGATGAGGGCTCGTATTTTCTGGTTT 507

QY 1115 acggggccccaggttcagcaagacaaagagctctctcatcagccggcgagttcctgg 1174

Db 506 ACGGGGCCCCAGGCTTCAGCAAAGACAAACGAGTCTCTCATCAGCCGGCGGAGTTCCTGG 447

QY 1175 ccggggtgcgggtcggggttccccaggttaagtacacctggcagccgaggtgtggtcctgc 1234

Db 446 CCGGGGTGCGGGTCCCGAGGTAACTGACCTGGCAGCCGAGGCTGTGGTCTCTGC 387

QY 1235 attacacagactggctgcctcccgaggagcccggcacgcctgagggagccctgagcgatg 1294

Db 386 ATTACACAGACTGGTGCATCCCGAGGACCCGGCACGCCTGAGGGAGGCCCTGAGCGATG 327

QY 1295 tgggtggcgaccacaatgctggtgcccccgctggccccagctggctggcgactggctgccc 1354

Db 326 TGGTGGCGGACCAACAATGTCGTGTGCNCCGTGGCCACAGCTGGCTGGCGACTGGCTGCC 267

QY 1355 agggtgccccggtctacgcctacgtctttgaacacccgtgcttccacgctctcctggcccc 1414

Db 266 AGGGTGCCGGGTCTACGCCTACGTCTTTGAACCCCGTGTCTCCACGCTCTCCTGGCCCC 207

QY 1415 tgtggtgggggtgccccacggctacgacgatcgagttcatcttgggatccccctggacc 1474

Db 206 TGTGGATGGGGTGCCCCACGGCTACGAGATCGAGTTTCATCTTTGGGATCCCCCTGGACC 147

QY 1475 cctctcgaactacacggcgagaggagaaaaatcttcgcccagcgactgatcgatactggg 1534

Db 146 CCTCTCGAAACTACACGGCAGAGGAGAGAAATCTTCGCCACGCGACTGATCGGATCTGGG 87

QY 1535 ccaactttcccgcacaggggatcccaatgatgccccgagaccccccaagaatggc 1594

Db 86 CCAACTTTGCCGCACAGGGGATCCCAATGAGCCCCGAGACCCCAAGGCCCCACAAATGGC 27

QY 1595 ccccgtaacggcgggggtcagcag 1620

Db 26 CCCCGTACACGGCGGGGCTCAGCAG 1

RESULT 13

BE466391/c

LOCUS

DEFINITION

BE466391

562 bp mRNA linear EST 27-JUL-2000

hz21b04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3208591 3' similar to gb:M55040 ACETYLCHOLINESTERASE PRECURSOR (HUMAN);;contains element MSR1 repetitive element ;, mRNA sequence.

BE466391

BE466391.1 GI:9512253

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 562)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 469.
FEATURES source
1. .562
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3208591"
/clone_lib="NCI_CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 104 a 184 c 173 g 97 t 4 others
ORIGIN
Query Match 32.3%; Score 558; DB 10; Length 562;
Best Local Similarity 99.3%; Pred. No. 8.4e-97;
Matches 558; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1066 caggtgctggtggtgaaggatgagggtcgtatttctgttttacggggcccca 1125
|||||
Db 562 CAGGTGCTGGTGGTGGAAGGATGAGGGCTCGTATTTCTGGTNTACGGGCCNCA 503
QY 1126 ggtctcagcaagacaacgagtctctcatcagccggccgagttcctggttcggcgggtgcgg 1185
|||||
Db 502 GCGTTCAGCANAGACAACGAGTCTCTCATCAGCCGGCCGAGTCTCTGCCGGGTGCGG 443
QY 1186 gtccgggttccccaggtaagtacctggcagccgaggtcgtggctcctgcattacacagac 1245
|||||
Db 442 GTCGGGGTCCNCAGGTAAGTGACCTGGCAGCCGAGGCTGTGCTCTGCATTACACAGAC 383
QY 1246 tggctgcatcccgaggaccggcacgcctgaggaggccctgagcgatgtggtggcgac 1305
|||||
Db 382 TGGCTGCATCCCGAGGACCCCGCAGCCTGAGGGAGGCCCTGAGCGATGTGGTGGCGGAC 323
QY 1306 cacaatgctggtgcccccgctggcggcgactggctgccccagggtgccccgg 1365
|||||
Db 322 CACAATGTGCTGTGCCCCCGTGCCAGCTGGCTGGCGACTGGCTGCCAGGGTGCCCCG 263
QY 1366 gtctacgcctacgtctttgaaacaccgtgtctccacgctctcctggccccctgtggatgggg 1425
|||||
Db 262 GTCTACGCCCTACGTCTTTGAACACCGTGTCTCCACGCTCTCTGGCCCCCTGTGGATGGG 203
QY 1426 gtgccccacggctacgagatcgagttcatctttgggataccccctggacccctctcgaaac 1485
|||||
Db 202 GTGCCCCACGGCTACGAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAAC 143
QY 1486 tacacggcagaggagaaaaatcttcgccccagcgactgatgcgatactgggccaactttgcc 1545
|||||
Db 142 TACACGGCAGAGGAGAAAATCTTCGCCACGCGACTGATCGGATACTGGGCCAACTTTGCC 83
QY 1546 cgcacaggggatcccaatgagccccgagaccccccaaggccccacaatggcccccgtaacg 1605
|||||

Db 82 CGCACAGGGGATCCCAATGAGCCCGAGACCCCAAGGCCCAACAATGGCCCCGTACACG 23
QY 1606 gcgggggtcagcagtagcttta 1627
|||||
Db 22 GCGGGGGCTCAGCAGTACGTTA 1
RESULT 14
BG395920 689 bp mRNA linear EST 12-MAR-2001
LOCUS 602458524F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4580980 5',
DEFINITION mRNA sequence.
ACCESSION BG395920
VERSION BG395920.1 GI:13289368
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 689)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM1300 row: p column: 05
High quality sequence stop: 685.
FEATURES source
1. .689
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4580980"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 94 a 234 c 223 g 138 t
ORIGIN
Query Match 32.2%; Score 555.2; DB 10; Length 689;
Best Local Similarity 98.1%; Pred. No. 3e-96;
Matches 604; Conservative 0; Mismatches 8; Indels 4; Gaps 4;
QY 1 atgagggccccgcagtgctgtgcacacgccttccctgggtcccaactccttcctc 60
|||||
Db 77 ATGAGGCCCCCGCAGTGTCTGTGCACACGCTTCCCTGGCTTCCCACTCCTTCTCCTC 136
QY 61 ctccctcggctcctcctgggtggaggagtgggggctgagggccggaggatgcagagctgctg 120
|||||
Db 137 CTCCTCTGGCTCCTGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 196
QY 121 gtgacggtgcgtggggccggcgtgcgggggcatcgcctgaagacccccggggccctgtc 180
|||||
Db 197 GTGACGGTGGTGGGGCCGGCTGCGGGGCATTCGCCCTGAAGACCCCGGGGCCCTGTC 256
QY 181 tctgcttctcctgggcatccccctttggggagccaccccatgggaccccgctcgtttctgcc 240
|||||
Db 257 TCTGCTTCTCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTTCTGCCA 316

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 15:34:03 ; Search time 153.1 Seconds
(without alignments)
2767.588 Million cell updates/sec

Title: US-09-810-861B-5
Perfect score: 1725
Sequence: 1 atgagggcccccgcagtgctct.....aattgctcagggccacctga 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1722.4	99.8	3096	2	US-08-318-826A-6
2	1722.4	99.8	3096	2	US-08-370-156-3
3	1722.4	99.8	3096	3	US-08-814-095-3
4	1722	99.8	1845	1	US-07-732-962A-1
5	1722	99.8	1845	5	PCT-US92-06106-1
6	1722	99.8	2256	2	US-08-318-826A-5
7	1722	99.8	2256	2	US-08-370-156-1
8	1722	99.8	2256	3	US-08-814-095-1
9	1722	99.8	3016	2	US-08-318-826A-7
10	1722	99.8	3016	2	US-08-370-156-5
11	1722	99.8	3016	3	US-08-814-095-5
12	1198	69.4	35060	3	US-08-814-095-7
13	380	22.0	2381	2	US-08-318-826A-9
14	380	22.0	2416	2	US-08-318-826A-8
15	380	22.0	2416	4	US-09-334-489-1
16	380	22.0	2416	4	US-09-334-489-2
17	378.4	21.9	2400	6	5215909-13
18	322.8	18.7	2445	6	5215909-9
19	185.6	10.8	764	6	5215909-7
20	142.6	8.3	3018	1	US-08-347-718B-3
21	142.6	8.3	3018	1	US-08-482-262-3
22	141	8.2	3018	6	5200183-1
23	139.4	8.1	2184	1	US-08-445-050-8
24	139.4	8.1	2184	1	US-08-204-691-8
25	139.4	8.1	2428	1	US-08-445-050-1
26	139.4	8.1	2428	1	US-08-204-691-1
27	139.4	8.1	2487	3	US-08-370-223-12

28	134.6	7.8	1907	1	US-08-462-884A-2	Sequence 2, Appl
29	134.6	7.8	1908	1	US-08-461-881B-2	Sequence 2, Appl
30	134.6	7.8	1908	2	US-09-123-960-2	Sequence 2, Appl
31	104.2	6.0	1753	4	US-09-058-260-21	Sequence 21, Appl
32	104.2	6.0	1756	4	US-09-058-260-31	Sequence 31, Appl
33	104.2	6.0	1776	4	US-09-058-260-23	Sequence 23, Appl
34	104.2	6.0	1896	4	US-09-058-260-3	Sequence 3, Appl
35	104.2	6.0	1952	4	US-09-058-260-5	Sequence 5, Appl
36	104.2	6.0	1957	4	US-09-058-260-19	Sequence 19, Appl
37	104.2	6.0	4090	2	US-08-781-802-5	Sequence 5, Appl
38	104.2	6.0	4090	4	US-08-694-078-5	Sequence 5, Appl
39	104.2	6.0	6263	2	US-08-781-802-3	Sequence 3, Appl
40	104.2	6.0	6263	4	US-08-694-078-3	Sequence 3, Appl
41	102.6	5.9	1699	4	US-09-058-260-13	Sequence 13, Appl
42	100.2	5.8	1800	1	US-08-484-815-11	Sequence 11, Appl
43	100.2	5.8	1800	3	US-08-888-949-11	Sequence 11, Appl
44	100.2	5.8	1800	4	US-08-888-950-11	Sequence 11, Appl
45	100.2	5.8	1800	4	US-09-262-758-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-318-826A-6
; Sequence 6, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
; OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,

OTHER INFORMATION: 3, 4, 5 and the translated portion of Intron 4 (readthrough)
US-08-318-826A-6

Query Match 99.8%; Score 1722.4; DB 2; Length 3096;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgaggccccgcagtgctgctgcacacgccttccctggcttccccactccttctctc 60
Db 160 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCACACTCCTTCTCTC 219
QY 61 ctctctggtcctcctggtgagagtggtgggctgagggccgggaggtgagagtgctg 120
Db 220 CTCTCTGGCTCCTGGGTGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279
QY 121 gtgacggtgctggtgggcccgcctgctgggcatctgcctgaagacccccggggccctgtc 180
Db 280 GTGACGGTGCCTGGGGCCGCTGCGGGCATTCGCCCTGAAGACCCCCGGGGCCCTGTC 339
QY 181 tctgcttctcctgggcatcccccttgcggagccaccctgggacccccgccttcttctgcca 240
Db 340 TCTGCTTCTCTGGGCATCCCTTTGCGGAGCCACCCATGGGACCCCTCGCTTCTGCCA 399
QY 241 ccggagcccaagcagccttggtcaggggtggtgtagacgtacacaccttccagagtgtctgc 300
Db 400 CCGGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacacccctataccccagggttttgagggcacccagagatgtggaacccaaac 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcaggtgctgctgacctaactcaacgtgtggagccacccataccccggcctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCTGTACCTCAACGTGTGGACACCATACCCCGGCCCTACA 579
QY 421 tccccacccctgctcctgctggtgatctatgggggtggcttctacagtggggcctcctcc 480
Db 580 TCCCCACCCCTGTCTCTGCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 639
QY 481 ttggacgtgtacgatggcgcttctgtgtacagggccgagagactgtgctggtgtccatg 540
Db 640 TTGGACGTGTACGATGGCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 699
QY 541 aactaccgggtgggagccttggcttccctggccctgcggggagccgagagggcccgggc 600
Db 700 AACTACCGGGTGGGAGCCTTTGGCTTCTTGGCCCTGCCGGGGAGCCGAGAGGGCCCGGGC 759
QY 601 aatgtgggtcctcctggtacagaggtggtccctgcagtggtggtgaggaacgtggcagcc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttcgggggtgaccccgacatcagtgacgtgtttggggagagcgcgggagccgctcggtg 720
Db 820 TCGGGGGGTGACCCGACATCAGTGACGCTGTTTGGGAGAGCGCGGGAGCCCGCTCGGTG 879
QY 721 ggcagtcacctgctgtcccgcccgccagccgggctgttccacagggccgtgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCGCCCGCAGCCGGGCCCTGTTCACAGGGCCGTGCTGCAGAGC 939
QY 781 ggtgcccccaatggaccctggggccacggtgggcatgggagagggccctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACCGTGGGCATGGGAGAGGCCCTCGCAGGGCCACG 999
QY 841 cagctggccccacctgtgggctgtcctccagggcgacactggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCCCACCTTGTGGGCTGTCTCCAGGGCGGCACTGGTGGGAATGACAGAGCTG 1059
QY 901 gtgacctgcttcggacacgaccagcaggtcctggtgaaccacgaatggcagctgctg 960
Db 1060 GTAGCCTGCCTTCGGACACGACACCGCAGGCTCCTGTTGAACCAACCAATGGCAGCTGCTG 1119
QY 961 cctcaagaaagcgtcttccggttctcctctgctgctgtggtagatggagacttctcagt 1020

Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCTCTTCGTGCCTGTGGTAGATGGAGACTTCTCAGT 1179
QY 1021 gacacccagagggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
Db 1180 GACACCCAGAGGGCCCTCATCAACCGGGAGACTTCCACGGCCTGCAGGTGCTGTGGGT 1239
QY 1081 gtggtgaaggtgagggctctgtatttctggtttacggggccccaggcttcagcaaaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTCTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1299
QY 1141 aacgagctctcatcagccggggccgaggttccctggccgggggtgcgggttcgccccag 1200
Db 1300 AACGAGTCTCTCATCAGCCGGGCCGAGTTCTCTGGCCGGGGTGGGGTTCGCCAG 1359
QY 1201 gtaagtacctggcagccgaggtgtgtctctgcattacacagactggtgcacatcccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGTGCATCCGAG 1419
QY 1261 gacccggcacgcctgagggagggccctgagcgtgtgtggggcgaccacaatgtcgtgtgc 1320
Db 1420 GACCCGGCACGCCCTGAGGGAGGCCCTGAGCGATGTGTGGGCGACCAATGTCTGTGTGC 1479
QY 1321 cccgtggccagctggtggcgactggtgccccagggtgccccgggtctctacgcctacgtc 1380
Db 1480 CCCGTGGCCAGCTGGCTGGCGACTGGCTGCCAGGGTGGCCGGGTCTACGCCCTAGCTC 1539
QY 1381 tttgaacacccgtgcttccacgctcctcctggccccctgtggatgggggtgccccacgggtac 1440
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTCTGGCCCTGTGGATGGGGTGGCCACGGGTAC 1599
QY 1441 gagatcgagttcatctttggggatccccctggacccctctcgaactacacggcagaggag 1500
Db 1600 GAGATCGAGTTCACTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaaattctgcggcagcagctgactgatactgggccaacttgcggcagggggtgagatccc 1560
Db 1660 AAAATCTTGCCTCAGCGACTGTATCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagaccccaaggcccccaaatggcccccgctacacggcggggctcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCGGTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtctggacctggcgccgctggaggtgcggggggggtgcgccccagggcctgc 1680
Db 1780 TACGTAGTCTGGACCTGGCGCGCTGGAGGTGCGGGGGGGGTGCGCGCCCCAGGCGCTGC 1839
QY 1681 gccttctggaaccgcttctcctcccccaaatgtgctcagcgccacctg 1724
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGCCACCCGG 1883

RESULT 2
US-08-370-156-3
; Sequence 3, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
US-08-370-156-3

Query Match 99.8%; Score 1722.4; DB 2; Length 3096;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgagggcccccagtgctgctgcacacgccttccctggcttccccactccttctcctc 60
Db 160 ATGAGGCCCCCGCAGTGCTGCTGCACACGCGCTTCCCTGGCTTCCCCACTCTTCTCCTC 219

QY 61 ctcctctggctcctggggtgaggagtgggggtgagggccgggaggatgcagagctgctg 120
Db 220 CTCCTCTGGCTCCTGGGTGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279

QY 121 gtgacgggtgcgtgggggcccgtgcggggcattcgccctgaagacccccggggccctgtc 180
Db 280 GTGACGGTGCGTGGGGCCGGCTGCGGGGCGATTCGCCCTGAAGACCCCCGGGGCCCTGTC 339

QY 181 tctgctttcctgggcattccccctttgcggagccaccatgggacccccctgccttctgccca 240
Db 340 TCTGCTTTCCCTGGGCATCCCCCTTTGCGGAGCCACCCCATGGACCCCGCTCGCTTCTGCCA 399

QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttccagagtgtctgc 300
Db 400 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGCTAGACGCTACAACCTTCCAGAGTGTCTGC 459

QY 301 taccaatatgtggacaccctatatcccaaggttttgagggcaccgagatgtggaaccccaac 360
Db 460 TACCAATATGTGGACACCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 519

QY 361 cgtgagctgagcaggactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
Db 520 CGTGAGCTGAGCGGAGGACTGCCTGTACCTCACGCTGTGGACACCATACCCCGGCCCTACA 579

QY 421 tccccaccacctgtcctcgtctggatctatgggggtggcttctacagtggggcctcctcc 480
Db 580 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCCTCTCC 639

QY 481 ttggacgtgtacgatggccgcttcttggtacagggccgagaggactgtgctggtgccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGCTGTCCATG 699

QY 541 aactaccgggtggagcctttggcttccctggccctgccggggagccgagagccccgggc 600
Db 700 AACTACCGGTTGGAGCCTTTGGCTTCCCTGGCCCTGCCGGGGAGCCGAGAGSCCCCGGCC 759

QY 601 aatgtgggtctcctggatcagaggctggccctgcagtgggtgcaggagaaactgggcagcc 660
Db 760 AATGTGGGTCTCCTTGGATCAGAGGCTGGCCCTTCGAGTGGGTGCAGGAGAACGTGGCAGCC 819

QY 661 ttcgggggtgacccgacatcagtgacgctgtttggggagagcgcgggagccgctcggtg 720
Db 820 TTGCGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGAGCCCTCGGTCG 879

QY 721 ggcattgcacctgctgtccccgcgccagccgggcccgtgttccacagggccgtgctgcagagc 780
Db 880 GGCAATGCACCTGCTGTCCCCGCCACGCGGGGCCCTTTCACACAGGGCCGTGCTGCAGAGC 939

QY 781 ggtgcccccaatggacocctgggccacggtgggcatgggagagggcccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCTGGGCCACGCGAGGTGGGCATGGGAGAGGCCCTCGCAGGCCACG 999

QY 841 cagctggccccacctgtgggctgtcctccagggcagctggtgggaatgacacagagctg 900
Db 1000 CAGCTGGCCCCACCTTGTGGGCTGTCTCTCCAGGGCGCACTGGTGGGAATGACACAGAGCTG 1059

QY 901 gtagcctgccttcgggacacgacgacgagcgaggtccttcgtgcctgtggtagatggagacctcctcagt 1020
Db 1060 GTAGCCTGCCTTCGGACACGACGACGCGAGGTCTCTGTGAACCAACGAATGGCACGCTGCTG 1119

QY 961 cctcaagaaagcgtcttccgggttctccttcgtgccttcgtgcctgtggtagatggagacctcctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCTCTCTTCGTGCCTTGGTAGATGGAGACTTCTCTCAGT 1179

QY 1021 gacacccccagaggccctcatcaacgcgggagagacttccacggccctgcaggtgctgggtgggt 1080
Db 1180 GACACCCACAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT 1239

QY 1081 gtggtgaaggatgagggtcgtattttctgtgttaccggggcccccaggtcttcagcaaaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCCAGGGCTTCAGCAAAGAC 1299

QY 1141 aacgagtctctcatcagcggggccgagttccttcctggcgggggtgcgggtgttcccccag 1200
Db 1300 AACGAGTCTCTCATCAGCGGGGCCGAGTTCTCTGGCCGGGTGCGGGTCCGSGTTCGCCAG 1359

QY 1201 gtaagtgacctggcagcggaggtggtggttcctgcattacacagactggctgcacatcccgag 1260
Db 1360 GTAAGTGACCTGGCAGCGCGAGGCTGTGTCCTCGATTACACAGACTGGCTGCATCCCGAG 1419

QY 1261 gacccggcacgcctgaggggagggccctgagcgtatggtggtggcgaccacaaatgctgtgtgc 1320
Db 1420 GACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGTGGGGGACCCACAATGTCGTGTGC 1479

QY 1321 cccgtggcccagctggtgggcgactggctgccaggtgcccggtgtccccgggtctacgcctacgtc 1380
Db 1480 CCCGTGGCCCCAGCTGGCTGGCGGACTGGCTGCCAGGGTGCCCGGGTCTACGCCCTACGTC 1539

QY 1381 tttgaacacccgtgcttccacgctctcctggcccttgcggccctgtggatgggggtgcccccaggtac 1440
Db 1540 TTTGAACACCCGTGCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGTGCCTCCACGGCTAC 1599

QY 1441 gagatcgagttcatctttgggatccccctggacccctctcgaaactacacggcagagggag 1500
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659

QY 1501 aaaaattctcgccccagcgactgatgcgatacttggggccaactttgcccgcacaggggatccc 1560
Db 1660 AAAATCTTCGCCCCAGCGACTGATGCGGATACTGGGGCCAACCTTTGCCCCGACAGGGGATCCC 1719

QY 1561 aatgagcccccgagaccccaaggcccccaaatggcccccgctacacggtgggggtcctcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCCAACAATGGCCCCCGCTACACGGCGGGGCTCAGCAG 1779

QY 1621 tacgttagtctggacctggcccgctggaggtgcgggggggtgctgcgcgccccagggcctgc 1680
Db 1780 TACGTTAGTCTGACCTGCGGCCCGCTGGAGGTGCGGGGGGGGCTGCGCGCCCCAGGCCTGC 1839

QY 1681 gccttctggaaccgcttctctcccccaaatgtgctcagcgccacctg 1724
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCCAAATTGTCTCAGCGGCCACCCGG 1883

RESULT 3
US-08-814-095-3
; Sequence 3, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced AChE
; DESCRIPTION: comprising exons 2, 3, 4 and 5 as well as the translated portion
; DESCRIPTION: of Intron 4 (readthrough)"
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..1959
US-08-814-095-3

Db	400		CCGAGCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTTACAAACCTTCCAGAGTGTCTGC	459
QY	301		taccaatatgtggacacacctataccccaggtttttagggcaccgagatgtggaacccaac	360
Db	460		TACCAATATGTGGACACCCCTATACCCAGGTTTGGGGCACCAGATGTGGAACCCCAAC	519
QY	361		cgtgagctgagcgaggactgcctgtacctaactcaacgtgtggacaccataccccccgacctaca	420
Db	520		CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGCTGTGACACCATACCCCGGCTTACA	579
QY	421		tcctccacccctgtcctgtggtgatctatgggggtggttcttacagtggggacctctcc	480
Db	580		TCCCCCACCCCTGTCTCTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTTCTCTCC	639
QY	481		ttggacgtgtacgatggcgcttcttggtacagggccgagagagactgtgtgttccatg	540
Db	640		TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG	699
QY	541		aactaccgggtggagaccttgggttcttccctggccctgcggggagcgcgagagcccgggc	600
Db	700		AACTACCGGGTGGAGCCTTGGCTTCTCTGGCCCTGCCGGGAGCCGAGAGGCCCGGGC	759
QY	601		aatgtgggtctcctggtatcagaggtgccttgcctgcagtggtgtcaggaacgtggagcc	660
Db	760		AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAACTGGCAGCC	819
QY	661		ttcgggggtgaccccgacatcagtgacgtgttggggagagcgcgggagccgctcgggtg	720
Db	820		TTCGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGACGGCGGGAGCCCTCGGTG	879
QY	721		ggcatgcacctgtgtccccgcccagcggggcctgttccacagggccgtgtgcagagc	780
Db	880		GGCATGCACCTGTGTCCCCGCCAGCCGGGGCTGTTCCACAGGGCCGTGTCAGAGC	939
QY	781		ggtgcccccaatggacctggggccacggtgggcatgggagagggccgtgcagggccacg	840
Db	940		GGTGCCCCCAATGGACCTGGGCCACGGTGGGCATGGGAGAGGCCCGCTCGCAGGGCCACG	999
QY	841		cagctggcccacctgtgtgggtgtcctccagcgccgactggtgggaatgacacagagctg	900
Db	1000		CAGCTGGCCCCACCTTGTGGGTGTCTCTCCAGGGGCACTGGTGGGAATGACACAGAGCTG	1059
QY	901		gtagcctgccttcggacacgaccagcgcaggtcctgttgaaccacgaatggcagctgctg	960
Db	1060		GTAGCCTGCCTTCGGACACGACCCAGCGAGGTCTGTGGTGAACCCACGAATGGCACGCTGTG	1119
QY	961		cctcaagaaagcgtcttcgggttctccttcctgcctgtgctgtgtagatggagaccttccagt	1020
Db	1120		CCTCAAGAAAGCGTCTTCCGGTCTCTCTTCGTGCCTGTGTGGTAGATGGAGACTTCTCTCAGT	1179
QY	1021		gacacccccagaggccctcatcaacgcgggagacttccacggcctgcaggtgctgtggtg	1080
Db	1180		GACACCCCAGAGGCCCTCATCAACGGGGAGACTTCCACGGCCTGCAGGTGCTGTGGGT	1239
QY	1081		gtggtgaaggatgagggtcgtattttctgtgtttacggggccccaggttcagcaagac	1140
Db	1240		GTGGTGAAGGATGAGGGCTCGTATTTCTGTTTACGGGGCCCCCAGGCTTTCAGCAAGAC	1299
QY	1141		aacgagtctctcatcagcggggcccaggttccctgcgggggtgcgggtgcgggttcccgag	1200
Db	1300		AACGAGTCTCTCATCAGCGGGCCGAGTTCTTCCCGGGGTGCGGGTTCGCCAG	1359
QY	1201		gtaagtacacctggcagcgaggctgtgttcctgcattacacagactggctgcatcccgag	1260
Db	1360		GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCTGCTATACACAGACTGGCTGCATCCCGAG	1419
QY	1261		gacccggcacgcctgagggagggccctgagcgtgtgtgtggcgaccacaatgtcgtgtgc	1320
Db	1420		GACCCGGCACGCCTGAGGGAGGCCCTGAGCGATGTGTGGGGGACCCACAATGTCTGTGTGC	1479
QY	1321		cccgtggcccagctggttggcgactggtgccccaggggtgccccgggtctacgcctacgctc	1380

Query Match 99.8%; Score 1722.4; DB 3; Length 3096;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1723; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	atgagggccccgcagtgctgctgcacacgccttccctggcttccccactcttctcctc	60
Db	160	ATGAGGCCCCCGCAGTGTCTGTGTCACACGCCCTTCCCTGGCTTCCCCACTCTCTCTCCTC	219
QY	61	ctcctctggctcctgggtggagagtggtgggctgagggccgggagggatgcagagctgctg	120
Db	220	CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG	279
QY	121	gtgacggtgctggtggggccggtgcggggcattgcctgaagacccccggggccctgtc	180
Db	280	GTGACGGTGGTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCCGGGGCCCTGTC	339
QY	181	tctgcttctcctggcctcccccttgcggagccaccatgggaccccgctcgttcttgcca	240
Db	340	TCTGCTTCTCTGGGATCCCTTTGCGGAGCCACCATGGGACCCCGTCGCTTCTGCTCA	399
QY	241	ccggagcccaagcagccttgggtcaggggtggtagcgtctacaaccttccagagtgctgc	300

Db 1480 CCCGTGGCCAGCTGGCTGGCGGAGCTGGTGCTGCCACAGGGTGCCCGGGTCTACGGCTACGTG 1539
QY 1381 ttgtgaacaccgtgtctccacgctctcctgcccctgtggatgggggtgccccacggctac 1440
Db 1540 TTTGAACACCGTGTCCACGCTCTCCTGGCCCCCTGTGGATGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcgagttcatctttgggatccccctggaccctctcgaaactacacgagaggag 1500
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaatcttcgcccagagactgatcgatactgggccaactttgcccgcacaggggatccc 1560
Db 1660 ANAATCTTCGCCACGACTGATCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagacccccaaagccccacaaatggcccccggtacacggcggtctcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCAAGCCCCCACAAATGGCCCCGTACACGGCGGGGGCTCAGCAG 1779
QY 1621 tacgttagctggacctgcggccgtggaggtgcggcggggggtgcgcgcgccagccctgc 1680
Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGCGGGGGGCTGCGCGCCCAAGGCTGC 1839
QY 1681 gccttctggaaccgttctctcccccaaatgtgctcagggccacctg 1724
Db 1840 GCCTTCTGGAACCGCTCCTCCCCAAATTGCTCAGCGCCACCGG 1883

RESULT 4

US-07-732-962A-1
; Sequence 1, Application US/07732962A
; Patent No. 5248604
; GENERAL INFORMATION:
; APPLICANT: Fischer, Meir
; TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
; TITLE OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,962A
; FILING DATE: 19910722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
US-07-732-962A-1

Query Match 99.8%; Score 1722; DB 1; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagggccccgcagtgctgtgcacacgccttcccctggcttccccactccttctcctc 60
Db 1 ATGAGGCCCCCGCAGTGTCTGTCACACGCCTTCCCTGCTTCCCCACTCCTTCTCCTC 60
QY 61 ctccctctggctccctgggtgagggagtggtgggctgagggccgggagggatgcagagctgctg 120
Db 61 CTCCTCTGGCTCCTGGGTGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 120
QY 121 gtgacgggtgcgtggggccggctgcggggcattcgccctgaagacccccggggccctgtc 180
Db 121 GTGACGGTGC GTGGGGCCGGCTGCGGGGCATTTCGCCTGAAGACCCCCGGGGCCCTGTC 180
QY 181 tctgctttcctgggcatacccctttgaggagccaccatgggacccccgtcgtttcttgcga 240
Db 181 TCTGCTTTCCTGGGCATCCCTTTGTCAGGGGTGGTAGACGCCTACAACCTTCCAGAGTGTCTGC 240
QY 241 cgggagcccaagcagccttggtcagggtggtgtagacgtctacaaccttccagagtgtctgc 300
Db 241 CCGGAGCCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCCTACAACCTTCCAGAGTGTCTGC 300
QY 301 taccaaatatgtggacacccctataccccagggttttgagggcaccgagatgtggaacccccaa 360
Db 301 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGGCACCGAGATGTGGAACCCCCAAC 360
QY 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggacaccataccccccggcctaca 420
Db 361 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATAACCCCCGGCCTACA 420
QY 421 tccccacccccgtcctcgtctggtgatctatgggggtggtctctacagtggggcctcctcc 480
Db 421 TCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 480
QY 481 ttggacgtgtacgatggccgctcttgggtacagggccgagaggactgtgctgggtgccatg 540
Db 481 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 540
QY 541 aactaccgggtgggagcctttggcttccctgcccctgcgggggagccgagagggccccgggc 600
Db 541 AACTACGGGTGGGAGCCCTTTGGCTTCTTGGCCCTGCCGGGAGCCGAGAGGCCCCCGGGC 600
QY 601 aatgtgggtctcctggatcagaggctggccctgagtcagtggtgcaggagaaacgtggcagcc 660
Db 601 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAAACGTGGCAGCC 660
QY 661 ttcgggggtgaccccgacatcagtgacgtgttggggagagcgcgggagccgcctcggtg 720
Db 661 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 720
QY 721 ggcatgcacctgctgtcccgcccgccagccggggcctgttccacagggccgtgctgcagagc 780
Db 721 GGCATGCACCTGCTGTCCCCGCCAGCCGGGGCCTGTTCACAGGGCCGTGCTGCAGAGC 780
QY 781 ggtgcccccaatggaccctggggccacggtgggcatgggagagggccccgtcgcagggccacg 840
Db 781 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCCGTTCGACGGGCCACG 840
QY 841 cagctggccccaccttgtgggtgtcctccaggcgccactgggtgggaatgacacagagctg 900
Db 841 CAGCTGGCCCCACCTTGTGGGTGTCTCTCCAGGGGGCACTGGTGGGAATGACACAGAGCTG 900
QY 901 gtgacctgccttcggacacacacagcgaggctcctgtgtgaaccacgaatggcacgtgctg 960
Db 901 GTAGCCTGCCTTCGGACACGACACGACGCGAGGTCTCTGGTGAACCCACGAATGGCACGCTG 960
QY 961 cctcaagaaagcgtcttccggttctccttctcgtgcctgtggtagatggagacttccctcagt 1020
Db 961 CCTCAAGAAAGCGTCTTCCGGTCTCTCCTTTCGTGCCTGTGTGTAGATGGAGACTTCCCTCAGT 1020
QY 1021 gacacccccagaggccctcatcaacgcgggagagacttccacggcctgcaggtgctgggtg 1080

Db 1021 GACACCCAGAGGCCCTCATCAACGGGAGACTTCCACGGCCTGCAGGTGCTGGGT 1080
QY 1081 gtggtgaagatgagggctgattttctgtttacggggcccccaggcttcagcaagac 1140
Db 1081 GTGGTGAAGGATGAGGGCTCGTATTTCTGGTTTACGGGGCCCCAGGCTTCAGCAAAGAC 1140
QY 1141 aacgagtctctcatcagccgggcccagatttccctggccggggtcggggttccccag 1200
Db 1141 AACGAGTCTCTCATCAGCCGGCCGAGTTCTCTGGCCGGGGTGGGGTTCCTCCAG 1200
QY 1201 gtaagtacctggcagccaggctgtgttctctgcattacacagactggctgcatcccgag 1260
Db 1201 GTAAGTGACCTGGCAGCCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCGAG 1260
QY 1261 gacccggcacgcctgagggagccctgagcgtgtgtgggagccacaaatgtcgtgtgc 1320
Db 1261 GACCCGGCACGCCCTGAGGGAGGCCCTGAGCGATGTGGTGGGCGACCCACAAATGCTGCTGC 1320
QY 1321 cccgtggccagctggctggcgactggctggccagggtgcccggtctacgcctacgtc 1380
Db 1321 CCCGTGGCCAGCTGGCTGGCGGACTGGCTGCCAGGGTGCCTGGTCTACGCCCTACGTC 1380
QY 1381 ttgaaacaccgtgttccacgctctcctggccctgtggtggggtggtgccccacggctac 1440
Db 1381 TTTGAACACCGTGTCTCCACGCTCTCTGGCCCTGTGGATGGGGTGCCCCACGGCTAC 1440
QY 1441 gagatcgagttcatctttgggatccccctggacccctctcgaactacacagcgagaggag 1500
Db 1441 GAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1500
QY 1501 aaaatcttcgcccagcgactgctgcatactgggcccactttgcccgcacaggggatccc 1560
Db 1501 AAAATCTTCGCCACGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1560
QY 1561 aatgagccccgagaccccaaggcccccaaatggcccccgatcacagcgcggtgctcagcag 1620
Db 1561 AATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1620
QY 1621 tacgttagtctgacctgcggccgctggaggtgcggggggtgcgcgccccaggtcgc 1680
Db 1621 TAGGTTAGTCTGGACCTGCGCCGCTGGAGGTGCGGGGGGGCTGCGCGCCAGGCGCTGC 1680
QY 1681 gccttctggaaccgttctctccccaaattgctcagcgccacc 1722
Db 1681 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGCCACC 1722

RESULT 5
PCT-US92-06106-1
; Sequence 1, Application PC/TUS9206106
; GENERAL INFORMATION:
; APPLICANT: Fischer, Melr
; TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
; TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, Esq.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06106
; FILING DATE: 19920722
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.
; REGISTRATION NUMBER: 28.678
; REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1845 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1842
PCT-US92-06106-1

Query Match 99.8%; Score 1722; DB 5; Length 1845;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atgagggcccccagtgctgtgctgcacacgccttccctggcttccccactctctctc 60
Db 1 ATGAGGGCCCCCGAGTGTCTGCTGCACACGCTTCCCTGGCTTCCCCACTCTCTCCTC 60
QY 61 ctctctgtgctcctgggtgagggagtggggctgagggcggggaggtgcagagctgtgc 120
Db 61 CTCCTCTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCGGGAGGATGCAGAGCTGCTG 120
QY 121 gtacgggtgcgtggggccggctgcggggcattgcctgaagacccccccggggccctgtc 180
Db 121 GTGACGGTGC GTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCCCGGGGCCCTGTC 180
QY 181 tctgctttcctgggcatccccctttgaggagccaccccatgggacccccgtcgtttctgcc 240
Db 181 TCTGCTTTCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTCTGCCA 240
QY 241 cgggagcccaagcagccttggtcagggggtggtagacgcctacaaaccttccagagtgtgc 300
Db 241 CCGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTTACAACTTCCAGAGTGTCTGC 300
QY 301 taccaaatatgtggacacccctataccccaggttttggagggccagagatgtggaaccccaac 360
Db 301 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAAC 360
QY 361 cgtgagctgagcagggactgcctgtacctacctaagctgtggacaccataccccccggcctaca 420
Db 361 CGTGAGCTGAGCGAGGACTGCTGTACTCTCAACGCTGTGGACACCATACCCCGGCCCTACA 420
QY 421 tccccacccctgtcctcgtctggatctatgggggtggtggtctctacagtggggcctctcc 480
Db 421 TCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 480
QY 481 ttggacgtgtacgatggcccttcttggtacagggccgagaggactgtgctggtgtccatg 540
Db 481 TTGGACGTGTACGATGGCCCGCTTCTTGGTACAGGCCCGAGAGGACTGTGCTGCTGCCATG 540
QY 541 aactaccgggtgggagcctttggcttccctggccctgcgggggagccgagagggccccgggc 600
Db 541 AACTACCGGGTGGGAGCCCTTTGGCTTCTCTGGCCCTGCCCGGGAGCCCGAGAGGCCCGGGC 600
QY 601 aatgtgggtctcctggtatcagaggctggccctgcagtgggtgcaggagaactgtggcagcc 660
Db 601 AATGTGGGTCTCTCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 660
QY 661 ttccgggggtgaccccgacatcagtgacgtgtttgtgggagagcgcgaggccgctcggtg 720
Db 661 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 720
QY 721 ggcatgcacctgctgtccccggcccgagccggggccctgttccacagggccgtgctgcagagc 780

Db 721 GGCATGCACCTGCTGTCCCGCCAGCCGGGCCCTGTTCCACAGGGCGGTGCTGCAGAGC 780
QY 781 ggtgcccccaatggaacctggtggccacagtggtggccatgggagagggccgtgcagggccacg 840
Db 781 GGTGCCCCCAATGAGACCTTGGGCCACGGTGGGCATGGGAGAGGCCCGTCGCAGGGCCACG 840
QY 841 cagctggccacacctgtggtgctctcctcagggcggcactggtgggaatgacacagagctg 900
Db 841 CAGCTGGCCACACCTTGTGGGCTGTCTCCAGCGGCACACTGGTGGGAATGACACAGAGCTG 900
QY 901 gtagcctgccttcggacacgacgacgaggtcctggtgacccacgaatggcacgtgctg 960
Db 901 GTAGCCTGCCTTCGGACACGACGACGCGCAGTCTGTTGTAACCAACGAATGGCAGCTGCTG 960
QY 961 cctcaagaagcgtcttcgggttctcctcctgctgctggtggtagatggagacttccctcagt 1020
Db 961 CCTCAAGAAAGCGTCTTCCGGTTCTCCTTCGTGCTGTGGTAGATGGAGACTTCTCTCAGT 1020
QY 1021 gacacccccagagggccctcatcaacgcgggagagacttccacggcctgcaggtgctggtgggt 1080
Db 1021 GACACCCCCAGAGGCCCTCATCAACGCGGGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT 1080
QY 1081 gtggtgaagatgagggctcgtatttctggtttacggggccccagggcttcagcaaaagac 1140
Db 1081 GTGGTGAAGATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCAGGCTTCAGCAAAAGAC 1140
QY 1141 aacgagtctctcatcagcggcgaggtcctggtcctgacacagactggctgcacatcccgag 1200
Db 1141 AACGAGTCTCTCATCAGCGGGCGGAGTCTCTGGCGGGGTGCGGGTTCGCCAG 1200
QY 1201 gtaagtacctggcagcggaggtggtgctgcattacacagactggctgcacatcccgag 1260
Db 1201 GTAAGTGACCTTGGCAGCGGAGGCTGTGTTCTGCTGATTTACACAGACTGCTGCATCCCGAG 1260
QY 1261 gacccggcagcctgagggagggccctgagcgtggtggtggtggtggtggtggtggtggtg 1320
Db 1261 GACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGGTGGCGGACCAACAATGTCGTGTGC 1320
QY 1321 cccgtggccacagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 1380
Db 1321 CCCGTGGCCACAGTGGCTGGCGGACTGGCTGCCAGGGTGCCGGGTCTACGCCCTACGTC 1380
QY 1381 ttgtaacacgtgcttccacgctctcctgccccctgtggtggtggtggtggtggtggtggtg 1440
Db 1381 TTTGAACACCGTCTTCCACGCTCTCTCTGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1440
QY 1441 gagatcgagttcatctttgggatacccccttgagccccctctcgaactacacggcagaggag 1500
Db 1441 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1500
QY 1501 aaaaatttcgccccagcagctgatgcgatactgggccaactttgccccgacaggggatccc 1560
Db 1501 AAAATCTTCGCCCCAGCGACTGATCGGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1560
QY 1561 aatgagccccagagaccccccaaggcccccaaatggcccccggtacacggcgggggtcagcag 1620
Db 1561 AATGAGCCCCGAGACCCCCAAGGCCCCACAAATGGCCCCCGGTACACGGCGGGGGCTCAGCAG 1620
QY 1621 tacgttagtctgagacctgcggccgctggaaggtgcggcggtgctgcggccccagggcctgc 1680
Db 1621 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGGGGGGTGCGGCCCCAGGCCTGC 1680
QY 1681 gccttctggaacgcgttccctcccccaaatggtcagcgccacc 1722
Db 1681 GCCTTCTGGAACCGCTTCTCTCCCCAAATGCTCAGCGCCACC 1722

RESULT 6

US-08-318-826A-5
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:

APPLICANT: Soreq, Hermona
APPLICANT: Zakut, Haim
APPLICANT: Eckstein, Fritz
TITLE OF INVENTION: Synthetic Antisense
TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
TITLE OF INVENTION: Containing Them
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 5891725thwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,826A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 2391.00001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5055
TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2256 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: /note= "Splice variant: Exons 1, 2,
OTHER INFORMATION: 3, 4 and 6"
US-08-318-826A-5

Query Match 99.8%; Score 1722; DB 2; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctgtgtgcacacgccttccctggcttccccactctctcctc 60
Db 160 ATGAGGGCCCCGCAGTGTCTGTGCACACGCGCTTCCCTGGCTTCCCCACTCCTTCTCTC 219
QY 61 ctctctgtgctcctgggtggagagtggtgggtggtggtggtggtggtggtggtggtggtg 120
Db 220 CTCCTTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGGCCGGGAGGATGCAGAGTGCTG 279
QY 121 gtgacggtgctggtggggcccggtgctgctgggtggtggtggtggtggtggtggtggtg 180
Db 280 GTGACGGTGCCTGGGGCCCGGCTGCGGGGCATTGCGCTGAAGACCCCCGGGGGCCCTGTC 339
QY 181 tctgcttctcctgggcatccccctttgctggagggcccccatgggacccccgtcgtcttctgcca 240
Db 340 TCTGCTTTCTCTGGGCATCCCCCTTTGCGGAGGCCACCCCATGGGACCCCGTCGCTTCTGCCA 399
QY 241 ccggagccccaaagcagccttggtcaggggtggtgagcgcgtacacaccttccagagtgtctgc 300
Db 400 CCGGAGCCCCAAGACGCTTGGTCAGGGGTGGTAGCGGTGGTAGCGCTACAACCTTCCAGAGTGCTGTC 459
QY 301 taccaatatgtggacacccctataccagggttttggggccacgagatgtggaaccccaac 360

Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTTTGAGGGCACCGAGATGTGGAAACCCCAAC 519

Qy 361 cgtgagctgagcagagactgcctgtacacctcaacgtgtggacaccatacccccgccctaca 420

Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTTGGACACCATACCCCGCGCTACA 579

Qy 421 tccccacccccctgtctcgtctggtatctatggtgggtggtctctacagtgggacctctcc 480

Db 580 .TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCCTCC 639

Qy 481 ttggacgtgtacgatggccgcttcttggtaacggccgagagactgtgctgtgtccatg 540

Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 699

Qy 541 aactaccgggtgggagcctttggcttccctggccctgccggggagccgagagagcccgccg 600

Db 700 AACTACCGGGTGGGAGCCTTGGCTTCTTGGCTTCCCTGCCCGGAGCCGAGAGGCCCGGGC 759

Qy 601 aatgtgggtctctcgatcagaggtggtggccctgcagtgggtgcaggagaaagtggcagcc 660

Db 760 AATGTGGTCTCCTGATCAGAGGCTGGCCCTGCAGTGGTGCAGGAGAACGTGGCAGCC 819

Qy 661 ttccgggggtgacccgacatcagtgacgctgtttggggagagcgcgggagccgctcggtg 720

Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 879

Qy 721 ggcattgcacctgctgtcccccgcagccagccggggtgttccacagggccgtgctgcagagc 780

Db 880 GGCATGCACCTGCTGTCCCCGCCACGCGCGGGCCGTGTTCCACAGGGCCGTGCTGCAGAGC 939

Qy 781 ggtgcccccaatggacctgggcccacggtgggagtggttccacagggccgtgagggccacg 840

Db 940 GGTGCCCCCAATGGACCTGGGCCACGCTGGGCATGGGAGAGGCCCGTGCAGGGCCACG 999

Qy 841 cagctggcccacctgtgggctgtcctccaggcggcactggtgggaatgacacagagctg 900

Db 1000 CAGCTGGCCCCACCTTGTGGGCTGTCTCCAGGGCGCACTGGTGGGAATGACACAGAGCTG 1059

Qy 901 gtagcctgccttcgacacagaccagcagcgaggtcctggtgaaccacgaatggcacgtgctg 960

Db 1060 GTAGCCTGCCTTCGGACACGACCAACGCGCAGGTCCTGTTGAACCAACGAATGGCAGCTGTG 1119

Qy 961 cctcaagaaaagcgtctccggttctccttccttcctgctgctgtgtagatggagacttccctcagt 1020

Db 1120 CCTCAAGAAAAGCGTCTCCGGTCTCTCCTTCGTGCTGTGGTAGATGGAGACTTCTCTCAGT 1179

Qy 1021 gacacccagagccctcatcaacgcgggagacttccacggcctgcagtgctggtgggt 1080

Db 1180 GACACCCAGAGGCCCTCATCAACGCGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT 1239

Qy 1081 gttgtgaaggatgagggtcgtattttctgttttacggggcccccaggcttcagcaaaagac 1140

Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTCTGTGTTTACGGGGCCCCAGGCTTCAGCAAAAGAC 1299

Qy 1141 aacgagtctctcatcagccggccggccgagttccttgccgggggtgcgggttccccag 1200

Db 1300 AACGAGTCTCTCATCAGCCGGCCGAGTTCTCTGCGGGGGTGCGGGTTCGGGTTCGCCAG 1359

Qy 1201 gtaagtgaacctggcagccgaggtgtggtcctgcattacacagactggctgcacatccccgag 1260

Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCCGAG 1419

Qy 1261 gacccggcacgcctgagggagggccctgagcgatggtggggcgaccacaatgtcgtgtgc 1320

Db 1420 GACCCGGCACGCCCTGAGGGAGGCCCTGAGCGATGTGGTGGCGACCAACAATGTCTGTGTGC 1479

Qy 1321 cccgtggccccagctggctggcgactggctgccaggggtgcccggttctacgcctacgtc 1380

Db 1480 CCCGTGGCCCCAGCTGGCTGGCGGACTGGCTGCCAGGGTGCCCGGGTCTACGCCCTACGTC 1539

Qy 1381 ttgaaacacgtgtcttccacgctctcctggccctgtggtatgggggtgccccacggctac 1440

Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTCTGGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599

Qy 1441 gagatcgagttcatctttgggatccccctggaccctctcgaactacacgcgagaggag 1500

Db 1600 GAGATCGAGTTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659

Qy 1501 aaaatcttcgcccagcgaactgatcgatactgggccaactttgccccgcacaggggatccc 1560

Db 1660 AAAATCTTCGCCCAGCGACTGATCGATACTGGGCCAACTTTGCCCGCACAGGGATCCC 1719

Qy 1561 aatgagccccgagacccccaaaggccccacaatggcccccgctacacggcggggtcagcag 1620

Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCACAATGGCCCCCGTACACGGGGGGGCTCAGCAG 1779

Qy 1621 tacgttagtctggaacctgcggccgctggaggtgcggggggtgcgcgcagggcctgc 1680

Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGCGGGGGCTGCGCGCCCAAGGCTGC 1839

Qy 1681 gccttctggaaccgcttctcccccaaatgtctcagcgcaccc 1722

Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATTGCTCAGCGCCACC 1881

RESULT 7

US-08-370-156-1

; Sequence 1, Application US/08370156

; Patent No. 5932780

; GENERAL INFORMATION:

; APPLICANT: Soreq, Hermona

; APPLICANT: zakut, Haim

; APPLICANT: Shani, Moshe

; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR

; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Reising, Ethington, Barnard & Perry

; STREET: P.O. Box 4390

; CITY: Troy

; STATE: Michigan

; COUNTRY: US

; ZIP: 48099

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/370,156

; FILING DATE:

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Kohn, Kenneth I.

; REGISTRATION NUMBER: 30,955

; REFERENCE/DOCKET NUMBER: P-307 (Mulford)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (810) 689-3500

; TELEFAX: (810) 689-4071

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2256 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-370-156-1

Query Match 99.8%; Score 1722; DB 2; Length 2256;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgaggcccccgagtgctgtgcacacgcctccctggcttccccactccttctcctc 60

Db 160 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCTTCCCTGGCTTCCCCACTCCTCTCCTC 219

QY 61 ctccctctggtcctcctggtgagagtggtgggctgagggccgggagagatgcagagctgctg 120
Db 220 CTCTCTGGCTCCTGGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279
QY 121 gtgacggtgctggggccggctgctggggcattgcctgaagacccccggggccctgtc 180
Db 280 GTGACGGTGCCTGGGGCCGGCTGCGGGGCATTCCGCTGAAGACCCCCCGGGGGCCTGTC 339
QY 181 tctgcttctcctgggcatcccttctgcggagccacccatgggacccctgcgtcttctgcca 240
Db 340 TCTGCTTTCTTGGGCATCCCTTTTGGGAGCCACCCATGGACCCCGTCGCTTTCTGCCA 399
QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaccttccagagtgtctgc 300
Db 400 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacacccctatatccagggttttggggccacccagagtgtggaaccccaac 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcagagactgcctgtacctcaacgttgtggacaccatacccccgccctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGCTGGACACCATAACCCCGGCTACA 579
QY 421 tccccacccctgtcctcgtctggtatctatgggggtggcctctacagtggggcctcctcc 480
Db 580 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCTCCTCC 639
QY 481 ttggacgtgtacgatggcgccttcttggtacagggccgagagactgtgctggtgtccatg 540
Db 640 TTGGACGTGTACGATGSCCGCTTCTTGGTACAGGGCCGAGAGGACTGTGCTGGTGTCCATG 699
QY 541 aactaccgggtgggagccttggcttccctgcccgtgcgggggagccgagagggcccgggc 600
Db 700 AACTACCGGGTGGGAGCCTTTGGCTTCCCTGGCCCTGCCGGGGAGCCGAGAGGCCCGGGC 759
QY 601 aatgtgggtcctcgtgatcagaggtgcccctgcaagtgggtgcaggagaacgtggcagcc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttcgggggtgacccgacatcagtgacgctgtttggggagagcgcgggagccgcctcggtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 879
QY 721 ggcattgcacctgctgtcccgcccgccagccggggcctgttccacagggccgtgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCCGCCCCAGCCGGGSCCTTCTCCACAGGGCCGTGCTGCAGAGC 939
QY 781 ggtgcccccaatggaccctggccacggtgggcattgggagagggcccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCTTGGCCACCGTGGGCATGGGAGAGGCCCTCGCAGGGCCACG 999
QY 841 cagctggccccaccttggtggcctgtcctccaggcgcactggtggaatgacacagagctg 900
Db 1000 CAGCTGGCCCCACCTTGTGGGCTGTCTCTCCAGCGGCACTGGTGGGAATGACACAGAGCTG 1059
QY 901 gtacgctgccttcggacacgacagcagcaggtctcctcgtgcctgtggtagatgggacgtcgtg 960
Db 1060 GTAGCCTGCCTTCGGACACGACGACGACGCTCCTCCAGCGGCACTTCCAGGCTGCAGTGTGCTG 1119
QY 961 cctcaagaaagcgtcttcgggttctcctcgtgcctcgtggtagatgggacgttctcagt 1020
Db 1120 CCTCAAGAAAGCGCTTCCGGTTCTCCTTCGTGCCTGTGGTAGATGGAGACTTCTCAGT 1179
QY 1021 gacacccagaggccctcatcaacgcgggagagacttccacggcctgcagggtgctggtgggt 1080
Db 1180 GACACCCACAGAGCCCTCATCAACCGGGGAGACTTCCAGGCTGCAGGTGCTGGTGGGT 1239
QY 1081 gtgggtgaaggatgagggctcgatttctgtgtttacggggggccccaggcttcagcaagac 1140
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGCCCCCAGGCTTCAGCAAGAC 1299
QY 1141 aacgagtctctcatcagccgggcccaggttccctggccgggggtgcgggtcgggttccccag 1200

Db 1300 AACGAGTCTCTCATCAGCCGGCCGAGTTCTCTGGCCGGGGTGGGGTCCCCAG 1359
QY 1201 gtaagtacctggcagccgagggctgtgctcctgcattacacagactggctgcatcccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCTTACACAGACTGGCTGCATCCCGAG 1419
QY 1261 gacccggcacgcctgagggagggccctgagcgtgtgtggggcgaccacaaatgctgtgtgc 1320
Db 1420 GACCCGGCACGCTGAGGGAGGCCCTGAGCGATGTGGTGGGCGACCACAAATGCTGTGTGC 1479
QY 1321 cccgtggccagctggctgggagcactggctggccagggtgcccgggtctacgcctacgtc 1380
Db 1480 CCCGTGGCCAGCTGGCTGGCGGACTGGCTGCCAGGSGTCCCGGGTCTACGCCCTACGTC 1539
QY 1381 ttgaaacaccgtgcttccacgctcctcctggccctgtggtgggtgccccacggctac 1440
Db 1540 TTTGAACACCGTGTCTCCACGCTCTCTGGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcgagtctcatttgggataccccctggacccctctgaaactacacggcagagag 1500
Db 1600 GAGATCGAGTTCTTTGGGATCCCCCTGGACCCCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaaatttcgcccagcagctgactgatacttgggccaacttggcccgacaggggacccc 1560
Db 1660 AAAATCTTCGCCCAGCGACTGATGCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagacccccaaaggcccccaaatggccccgtacacggcggggctcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCCAAGCCCCCACAATGGCCCCCTACACGGGGGGCTCAGCAG 1779
QY 1621 tacgttagtctggacctgcggccgctggaggtgcggcgggggtgcgcgccccagggcctgc 1680
Db 1780 TACGTTAGTCTGGACCTGCGCCCGCTGGAGGTGCGGGGGGGCTGCGCGGCCCTGTC 1839
QY 1681 gccttctggaacggcttctcccccacaaattgctcagcgccacc 1722
Db 1840 GCCTTCTGGAACCGCTTCTTCCCCAAATTTGCTCAGCGGCCACC 1881

RESULT 8

US-08-814-095-1
; Sequence 1, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; City: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050

TELEFAX: (248) 539-5055
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2256 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "ACHE gene comprising exons
DESCRIPTION: 2, 3, 4 and 6"
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-814-095-1

Query Match 99.8%; Score 1722; DB 3; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctgtgcacacgccttcctggcttccccactccttctc 60
Db 160 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCCACTCTCTCCTC 219
QY 61 ctctcttgctcctgggtggagagtggtgggctgagggccgggaggtgcagagctgtg 120
Db 220 CTCCTCTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGTG 279
QY 121 gtgacggtgcgtggggccggtgcggggcattcgccctaagacccccggggccctgtc 180
Db 280 GTGACGGTGCCTGGGGCCGGCTGCGGGGCATTTCGCCCTGAAGACCCCCGGGGCCCTGTC 339
QY 181 tctgctttcctgggcattcccctttgcggagccaccaccatgggacccccgcgtcttctgcc 240
Db 340 TCTGCTTTCCTGGGCATCCCCCTTTCGGGAGCCACCCATGGGACCCCCCGTCTGCTTCTG 399
QY 241 cggagcccaagcagccttgctcaggggtggtagacgctacacaccttccagagtgtctgc 300
Db 400 CCGAGCCCCAAGCAGCCCTTGGTCAGGGGTGGTAGACGCTACAACTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacaccctataccccaggtttgagggcaccgagatgtggaaccccaac 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTTCAGGGCACCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcaggaactgcctgtacctcaacgtgtggacaccatacccccgccctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCATATCCCCGGCCTACA 579
QY 421 tccccaccccctgtcctcgtctggatctatgggggtggtcttctacagtggggcctcctcc 480
Db 580 TCCCCACCCCTGTCTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCTCTCTCC 639
QY 481 ttggacgtgtacgatggccgcttctgtgtacagggccgagaggactgtgctggtgtccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 699
QY 541 aactaccgggtggagcctttggcttccctggccctgcgggggagccgagagggcccgggc 600
Db 700 AACTACCGGGTGGGAGCCCTTGGCTTCTTGGCCCTGCGGGGAGCCGAGAGGCCCGGGC 759
QY 601 aatgtgggtcctcgtgatacagaggctggccctgcagtggtgtcagagaaactgtggcagcc 660
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttggggggtgacccgacatcagtgacgtgctgtttggggagagcgcgggagccgctcgggtg 720
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 879
QY 721 ggcatacacctgctgtccccgcagccggggcctgtttccacagggccgtgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCCCGCCAGCCGGGGCCTGTTCACACAGGGCCGTGCTGCAGAGC 939

QY 781 ggtgcccccaatggaccctgggcccaggtgggcatgggagagggcccgctgcaggggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCGCTCGCAGGGCCACG 999
QY 841 cagctggccccaccttgtgggtgtcctccaggcggcactggtggaatgacacagagctg 900
Db 1000 CAGTGGCCCCACCTTGTGGGCTGTCTCCAGCGGCACCTGGTGGGAATGACACAGAGCTG 1059
QY 901 gtagctgccttcggacacgaccagcgaggtcctggtgaaccacgaatggcacgtgctg 960
Db 1060 GTAGCCTGCCTTCGGACACGACCGCAGCGAGGTCTGTGTGAACACGAATGGCACGTGCTG 1119
QY 961 cctcaagaaagcgtcttccgggttctccttctcctgtgcctgtggtagatggagacttccctcagt 1020
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCCTTCTCGTGCCTGTGTTAGATGGAGACTTCCTCTCAGT 1179
QY 1021 gacaccccgagaggccctcatcaacgcgggagacttccacggcctgcaggtgctggtgggt 1080
Db 1180 GACACCCACAGAGGCCCTCATCAACCGGGGAGACTTCCACGGCTGCAGGTGCTGGTGGGT 1239
QY 1081 gtgtgaaggatgagggctcgatatttctgttttacggggccccaggcttcagcaaaagac 1140
Db 1240 GTGTTGAAGGATGAGGGCTCGTATTTCGTGTTTACGGGGCCCCAGGCTTCAGCAAAAGAC 1299
QY 1141 aacgagtctctcatcagccggcgaggttccctggccgggtgctgggttcccccag 1200
Db 1300 AACGAGTCTCTCATCAGCCGGGCGAGTTCCTGGCCGGGTGCGGGTTCGCCCAG 1359
QY 1201 gtaagtacacctggcagccgaggtgtggtcctgcattacacagactggctgcacccccag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCCGAG 1419
QY 1261 gaccgggcacgcctgagggagggccctgagcgtggtggtggcgaccacaatgtcgtgtgc 1320
Db 1420 GACCCGGCACGCCCTGAGGGAGGCCCTGAGCGATGTGTTGSGCGACCAATAATGTCTGTGTC 1479
QY 1321 ccgctggccccagctggctggcgactggctgccccagggtgccccgggtctacgcctacgtc 1380
Db 1480 CCGTGGCCCCAGCTGGCTGGCGACTGGCTGCCACAGGGTGGCCGGGTCTACGCCCTACGTC 1539
QY 1381 ttgtaacacccgtgtcttccacgctctcctggccctgccccctgtggtgggggtgccccacggctac 1440
Db 1540 TTGTAAACACCGTGTCTCCACGCTCTCTGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcgagttcatctttgggataccccctggacccctctgaaaccttggccacaggggatccc 1500
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaaatctcgccccagcagctgatacgatactggggccaaacttggccacaggggatccc 1560
Db 1660 AAAATCTTCGCCCAGCGACTGATCGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagaccccccaaggcccccaaatggcccccgctacacggcggggctcagcag 1620
Db 1720 AATGAGCCCCGAGACCCCCAAGCCCCCACAAATGGCCCCCTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtcttggaacctgcggccgctggaggtgcggggggggtgcgcgccccagggcctgc 1680
Db 1780 TAGCTTAGTCTGGACCTGGGCGCTGGAGGTGCGGGGGGGGTGCGCGGCCACAGGCTGC 1839
QY 1681 gccttctggaaccgcttctcccccaaatgtctcagcgcaccc 1722
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCCAAATTTGCTCAGCGGCCACC 1881

RESULT 9
US-08-318-826A-7
; Sequence 7, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz

;
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,826A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 2391.00001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; OTHER INFORMATION: /note= "Splice Variant: Exons 1, 2,
; OTHER INFORMATION: 3, 4, 5 and 6"
; US-08-318-826A-7

Query Match 99.8%; Score 1722; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaggccccgcagtgctgctgcacacgccttcctcggttcctcccaactccttcctc 60
Db 160 ATGAGGCCCCCGCAGTGCTGCTGCACACGCCCTTCCCTGGCTTCCCCACTCCTTCTCCTC 219

QY 61 ctccctctggctcctgggtgagagagtggggctgagggcgaggagatgcagagctgctg 120
Db 220 CTCCTCTGGCTCCTGGGTGGAGGAGTGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTG 279

QY 121 gtgacggtgcgtggggcccggtcgcggggcattcgcctgaagaccccccggggcccctgtc 180
Db 280 GTGACGGTGC GTGGGGCCCGCTGCGGGGCATTCCGCTTGAAGACCCCCGGGGCCCTGTC 339

QY 181 tctgcttccctgggcatacccccttgcggagccacccatgggacccccgcgtcttctgcc 240
Db 340 TCTGCTTTCTTGGGCATCCCTTTGCGGAGCCACCCATGGGACCCCGTCGCTTTCTGCCA 399

QY 241 ccggagcccaagcagccttggtcaggggtggtagacgctacaaaccttcagagtgctgc 300
Db 400 CCGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTTCAAACCTTCCAGAGTGTCTGC 459

QY 301 taccaatatgtggacacccctataccagggttttgagggcacccgagatgtggaacccca 360
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGGCACCCGAGATGTGGNAACCCCA 519

QY 361 cgtgagctgagcgaggactgcctgtacctcaactcaacgtgtggacaccatacccccgccctaca 420
Db 520 CGTGAGCTGAGCGAGGACTGCCCTGTACCTCAACGTGTGGACACCATACCCCCGCCTACA 579

QY 421 tccccacccctgtcctcgtctggatctatgggggtggtctctacagtggggcctcctcc 480
Db 580 TCCCCACCCCTGTCCCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCTCCTCC 639

QY 481 ttggacgtgtacgatggccgctctcttggttacaggccgagagactgtgctggtgccatg 540
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGGTGTCCATG 699

QY 541 aactaccgggtgggagcccttggcttcctcctgcccctgcggggagccgagagccccgggc 600
Db 700 AACTACGGGTGGGAGCCCTTTGGCTTCTTGCCCTGCCGGGAGCCGAGAGGCCCGGGGC 759

QY 601 aatgtgggtctcctggatcagaggctggccctgcagtggtgcaggagaaacgtgcagcc 660
Db 760 AATGTGGGTCTCTTGGATCAGAGCTGGCCCTGCAGTGGTGGTGCAGGAGAAACGTGGCAGCC 819

QY 661 ttccgggggtgacccgacatcagtgcctgtttggggagagcgggagccgcctcggtg 720
Db 820 TTCCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGGGGAGCCGCTCGGTG 879

QY 721 ggcattgcacctgctgtcccccgcacgcccggggcctgttccacaggccgctgctgcagagc 780
Db 880 GGCATGCACCTGCTGTCCCCGCCACGCCGGGCCCTGTTCCACAGGGCCGCTGCTGCAGAGC 939

QY 781 ggtgcccccaaatggaccctgggccacaggtgggcatgggagaggccccgctcgcaggggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGGTGGGCATGGGAGAGGCCCGCTCGCAGGGGCCACG 999

QY 841 cagctggccccaccttggtgggtgtcctccagcggcactggtgcggaatgacacagagctg 900
Db 1000 CAGCTGGCCCCACCTTGTGGGCTGTCTCTCCAGGGGCACACTGGTGGGAATGACACAGAGCTG 1059

QY 901 gtagcctgccttcggacacgacagcagcaggtcctggtgaaccaacgaatggcacgtgctg 960
Db 1060 GTAGCTGCCCTTCGGACACGACGACGCGCAGGTCTCTGGTGAACCAACGAAATGGCACGTGCTG 1119

QY 961 cctcaagaaaagcgtcttcctcggttctccttcctgcctgctggtgtagatggagacttccctcagt 1020
Db 1120 CCTCAAGAAAAGCGTCTTCCGGTTCTCCTTCGTGCCCTGTGTGATGGAGACTTCCCTCAGT 1179

QY 1021 gacaccccagaggccccctcatcaacgcgggagagacttccacaggcctgcaggtgctggtgggt 1080
Db 1180 GACACCCACAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGGTGGGT 1239

QY 1081 gtggtgaaggatgagggtcgtattttctggttttacggggccccaggcttcagcaaaagac 1140
Db 1240 GTGTTGAAGGATGAGGGCTCGTATTCTGTTTACGGGGCCCCAGGCTTCAGCAAAGAC 1299

QY 1141 aacagagtctctcatcagccgggcccagagttcctgcccggggtgcgggtcccccag 1200
Db 1300 AACGAGTCTCTCATCAGCCGGGCGGAGTTCTTGCCGGGTGCGGGTTCGGGTTCCTCCAG 1359

QY 1201 gtaagtgcacctggcagccgaggctgtggtcctgcattacacagactggctgcatccccgag 1260
Db 1360 GTAAGTGACCTGGCAGCCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCGAG 1419

QY 1261 gaccggcacgcctgaggaggccccctgagcgtggtggggagaccacaatgctgctgtgc 1320
Db 1420 GACCCGGCACGCTTGAGGGAGSGCCCTGAGCGGATGTGGTGGGGCACCACAAATGTCGTGTGC 1479

QY 1321 ccggtggccccagctggctggggagactggctgcccagggtgccccgggtctacgcctacgtc 1380
Db 1480 CCGGTGGCCCCAGCTGGCTGGGGCAGCTGGCTGCCCCAGGGTGCCTGGGGTCTACGCCCTACGTC 1539

QY 1381 ttgaaacacccgtgcttccacgctctcctgccccctgtggatgggggtgccccacgggtac 1440
Db 1540 TTGAAACACCCGTGCTTCCACGCTCTCCTTGCCCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599

QY 1441 gagatcgagttcatctttgggatccccctggacccctctcgaatactacacggcagaggag 1500
|||||
Db 1600 GAGATCGAGTTCATCTTTGGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659
QY 1501 aaaatcttgcggccagcgactgatcgatacttgggccaaactttgcccgcacaggggataccc 1560
|||||
Db 1660 AAAATCTTCGCCCAGCGACTGATCGGATACTGGGCCAACTTTGCCCGCACAGGGGATCCC 1719
QY 1561 aatgagcccgagacccccaaagcccaaatgccccctgatacagggcggtcagcag 1620
|||||
Db 1720 AATGAGCCCGAGACCCCAAGCCCCACAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtgtgacctgcggccgctgaggtgcggcggtgctgcgcgccagggcctgc 1680
|||||
Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGGGGGCTGCGGCCCCAGGCCCTGC 1839
QY 1681 gccttctggaaccgcttcctcccccgaattgctcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCCTCCCCCAATTGCTCAGCGGCCACC 1881

RESULT 10
US-08-370-156-5
; Sequence 5, Application US/08370156
; Patent No. 5932780
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTICHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P.O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370.156
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-307 (Mulford)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
; US-08-370-156-5

Query Match 99.8%; Score 1722; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagccccgcagtgctgtgcacacgccttccctggcttccccactccttctcctc 60
|||||

Db 160 ATGAGGCCCCCGCAGTGTCTGTGCACACACGCTTCCCTGGCTTCCCCACTCCTTCTCCTC 219
QY 61 ctcccttggtcctcctgggtggaggagtgggggctgagggccgggagagatgcagagctgctg 120
|||||
Db 220 CTCCTCTGGCTCCTGGGTGGAGAGTGGGGGCTAGGGCCCGGAGATGCAGAGCTGCTG 279
QY 121 gtgacggtgctggggggccggtgcggtggggaattcgcctgaagacccccgggggcccctgtc 180
|||||
Db 280 GTGACGGTGCCTGGGGGCCCGGTGCGGGGCATTGCGCTGAAGACCCCGGGGGCCCTGTC 339
QY 181 tctgtttcttggtgcatccccctttggtgagccaccccatgggacccgtcgtttcttctgcc 240
|||||
Db 340 TCTGCTTTCCTGGGCATCCCCCTTTGCGGAGCCACCCATGGGACCCCGCTGCTTCTGCCA 399
QY 241 ccggagcccaagcagccttggtcaggggtggttagacgctacaaccttccagagtgtctgc 300
|||||
Db 400 CCGGAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGC 459
QY 301 taccaatatgtggacacccctatacccagggtttttagggcaccgagatgtggaaccccaac 360
|||||
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGAGGGCACCCGAGATGTGGAACCCCAAC 519
QY 361 cgtgagctgagcgaggactgcctgtacctcaacgtgtggaacaccataccccccggcctaca 420
|||||
Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACCTCAACGTGTGGACACCAATACCCCGGCTTACA 579
QY 421 tccccacccctgtcctcgtcgtggtatgtatgggggtggttctacagtggtgtctccatg 480
|||||
Db 580 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGGTGGCTTCTACAGTGGGGCCTCCTCC 639
QY 481 ttggacgtgtacgatggccgcttcttgggtacagggccgagagactgtgctggtgtccatg 540
|||||
Db 640 TTGGACGCTGACGATGGCCGCTTCTTGGTACAGGCCGAGAGGACTGTGCTGTTGTCCATG 699
QY 541 aactaccgggtgggagcctttggcttctcctggcctgcctgcggggagcggagagccccgggc 600
|||||
Db 700 AACTACCGGGTGGGAGCCTTTGGCTTCTGGCCCTGCCCGGGAGCCGAGAGCCCCCGGGC 759
QY 601 aatgtgggtctcctcgtgatacagaggtggtccctgcagtggtgcaggagaaactgtggcagcc 660
|||||
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819
QY 661 ttccggggtgaccccgacatcagtgacgtgttggggagagcggggagccgctcgtggtg 720
|||||
Db 820 TTCGGGGGTGACCCGACATCAGTGACGTGTTTGGGAGAGCGCGGAGCGCCCTCGGTG 879
QY 721 ggcatagcacctgctgtccccgcggccagcggggcctgttccacagggccgtgctgcagagc 780
|||||
Db 880 GGCATGCACCTGCTGTCCCCGCCACAGCCGGGGCCTGTTCACAGGGCCGTGCTGCAGAGC 939
QY 781 ggtgcccccaatggaccctgggccacggtgggcatgggagagggccgctgcagggccacg 840
Db 940 GGTGCCCCCAATGGACCCCTGGGCCACGTTGGGTCATGGGAGAGGCCCGCTGCGAGGGCCACG 999
QY 841 cagctggccccaccttgtgggctgtcctccaggcggcactggtgggaatgacacagagctg 900
|||||
Db 1000 CAGCTGGCCACCTTGTGGGCTGTCTCTCCAGGGCGCCTGTTGGGAATGACACAGAGCTG 1059
QY 901 gtacgtgccttcggacacgacacgagcaggttcctggtgaaccagaatggcacgtgctg 960
|||||
Db 1060 GTAGCCTGCCTTCGGACACGACGACGAGGTCTCTGGTGAACCAACGAATGGCACGTGCTG 1119
QY 961 cctcaagaaaagcgtcttccgggttctcctcgtgctgctggttagatggagacttcctcagt 1020
|||||
Db 1120 CCTCAAGAAAGCGCTTTCGGGTTCTCTCTGCTGCTGTGGTAGATGGAGACTTCTCTCAGT 1179
QY 1021 gacacccagagagccctcatcaacgggggagacttccacggcctgcaggtgctggtgggt 1080
|||||
Db 1180 GACACCCAGAGGCCCTCATCAACGGCGGAGACTTCCACGGCCCTGCAGGTGCTGGTGGGT 1239
QY 1081 gtggtgaaggatgagggctcgattttcttggttttacggggggccccaggcttcagcaaaagac 1140
|||||
Db 1240 GTGGTGAAGGATGAGGGCTCGTATTTTCTGGTTTACGGGGGGCCCCCAGGCTTCAGCAAGAC 1299

QY 1141 aacgagtgctctcatcagccgggcccagtgctcctggccgggggtgcggggtcggggttccccag 1200
|||||
Db 1300 AACGAGTCTCTCATCAGCCGGGCCGAGTTCTTGCCGGGGTGCGGGTGCGGGTTCCCCAG 1359

QY 1201 gtaagtgcctggcagcgaggctgtggtcctgcattacacagactggctgcacccgag 1260
|||||
Db 1360 GTAAGTGACCTGGCAGCGAGGCTGTGGTCTCTGCATTACACAGACTGGCTGCATCCCGAG 1419

QY 1261 gacccggcacgcctgaggagccctgagcgatgtggtggcgaccacaatgtcgtgtgc 1320
|||||
Db 1420 GACCCGGCACGCCTGAGGAGGCCCTGAGCGATGTGTTGGCGGACCAACAATGTCTGTGC 1479

QY 1321 cccgtggcccagctggctgggcgactggctgccccaggggtgcccgggtctctacgcctacgtc 1380
|||||
Db 1480 CCCGTGGCCACGCTGGCTGGCGACTGGCTGCCACGGGTGCCCGGTCTACGCCCTACGTC 1539

QY 1381 ttggaacacccgtgtctccacgctctcctcctggcccctgtggatgggggtgccccacgggtac 1440
|||||
Db 1540 TTTGAACACCGTGCTTCCACGCTCTCCTGGCCCCCTGTGGATGGGGTGTCCCCACGGCTAC 1599

QY 1441 gagatcgagttcatctttgggataccccctggaccctctctcgaactttgccccacagggatccc 1560
|||||
Db 1600 AAAATCTTGCCCACGCACTGATCGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1659

QY 1501 aaaatcttgcgccagcgactgatcgatactgggccaactttgccccacagggatccc 1560
|||||
Db 1660 AAAATCTTGCCCACGCACTGATCGGATCCCCCTGGACCCCTCTCGAAACTACACGGCAGAGGAG 1719

QY 1561 aatgagccccgagacccccaaaggccccacaaatggcccccgctacacggcgggctcagcag 1620
|||||
Db 1720 AATGAGCCCCGAGACCCCAAGGCCCCACAAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779

QY 1621 tacgttagtctggacctgcggccgctggaaggtgcgggggggtgcgcgcccaagcctgc 1680
|||||
Db 1780 TAGGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGGGGGGGGTGCGCGCCCAAGGCTGC 1839

QY 1681 gccttctggaaccgttctctccccaaaattgctcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCCAAATTGCTCAGCGCCACC 1881

RESULT 11
US-08-814-095-5
; Sequence 5, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3016 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Alternatively spliced AChE
; DESCRIPTION: comprising exons 2, 3, 4, 5 and 6"
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 160..2010
US-08-814-095-5

Query Match 99.8%; Score 1722; DB 3; Length 3016;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgagggccccgcagtgctctgcacacgccttccctggcttccccactcttctctc 60
|||||
Db 160 ATGAGGCCCCCGCAGTGTCTGCTGCACACGCCTTCCCTGGCTTCCCCACTCCTTCTCCTC 219

QY 61 ctctctggctcctgggtggaggagtggtgggctgagggcgaggagatgcagagctgctg 120
|||||
Db 220 CTCTCTGGCTCCTGGGTGGAGAGTGGGGCTGAGGGCCGGAGGATGCAGAGCTGCTG 279

QY 121 gtgacggtgcgtggggccggtgcggggcattcgccctgaagacccccggggccctgctc 180
|||||
Db 280 GTGACGGTGCTGTTGGGGCCGCTGCGGGGCATTTCGCCCTGAAGACCCCCGGGGCCCTGTC 339

QY 181 tctgcttctcctgggcatacccccttgcggagcaccacccatgggacccccgcgtcttctgcc 240
|||||
Db 340 TCTGCTTCTCTGGCATCCCCCTTTCGGGAGCCACCCATGGGACCCCGTCGCTTTCTGCCA 399

QY 241 ccggagcccaagcagccttgctcagggggtggtagacgctacaacctccagagtgtctgc 300
|||||
Db 400 CCGGAGCCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTTACAACCTTCCAGAGTGTCTGC 459

QY 301 taccaatatgtggacacctataccccaggttttgagggcacagagatgtggaaccccaac 360
|||||
Db 460 TACCAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGAACCCCAAC 519

QY 361 cgtgagctgagcgaggactgcctgtacctcaactcaactgtggacaccataccccggcctaca 420
|||||
Db 520 CGTGAGCTGAGCGAGGACTGCCTGTACTCTCAACGTGTGGACACCATACCCCGCCCTACA 579

QY 421 tccccacccccctgtcctcgtctggatctatgggggtggttctctacagtgggcctcctcc 480
|||||
Db 580 TCCCCACCCCTGTCTCGTCTGGATCTATGGGGTGGCTTCTACAGTGGGGCCCTCCTCC 639

QY 481 ttgacggtgtacgatggccgcttcttgggtacagggccgagagagactgtgctggttccatg 540
|||||
Db 640 TTGGACGTGTACGATGGCCGCTTCTTGGTACAGGGCCGAGAGACTGTGCTGGTGTCCATG 699

QY 541 aactaccgggtgggagccttggcttccctggcctgcctgcagtgccggagccgagagccccgggc 600
|||||
Db 700 AACTACCGGGTGGAGCCCTTGGCTTCTTGCCCTGCCGGGAGCCGAGAGCCCCGGGC 759

QY 601 aatgtgggtcctcctggatcagagggctggccctgcagtgagtggtgcagagagaacgtggcagcc 660
|||||
Db 760 AATGTGGGTCTCCTGGATCAGAGGCTGGCCCTTGCAGTGGGTGCAGGAGAACGTGGCAGCC 819

QY 661 ttcgggggtgacccgacatcagtgacgctgttttggggagagcgcgaggccctcggtg 720
|||||
Db 820 TTCGGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGGAGCCGCTCGGTG 879

QY 721 ggcattgcacctgctgttccccgccagccggggtgttccacagggccgtgtgctgcagagc 780

|||||
Db 880 GGCATGCACCTGCTGTCCCGCCAGCGGGGCTGTTCACACAGGGCGGTGTCAGAGC 939
QY 781 ggtgcccccaatggaacctgggccccaggtggccatgggagagggcccgctgcagggccacg 840
|||||
Db 940 GGTGCCCCCAATGAGACCTTGGGCCACGGTGGGCATGGGAGAGGCCCGTCGCAGGGCCACG 999
QY 841 cagctggccccactgtggctgtctctccagcgccactggcgccactgtgggaatgacacagagctg 900
|||||
Db 1000 CAGCTGGCCACCTTGTGGGCTGTCTCCAGCGGCCTGTGGTGAACCAACGAATGGCAGCTG 1059
QY 901 gtagcctgccttcggacacgaccagcgagctcctggtggaaccacgaatggcacgtgctg 960
|||||
Db 1060 GTAGCCTGCCTTCGGACACGACGACGCGCAGGTCCTGGTGAACCAACGAATGGCAGCTG 1119
QY 961 cctcaagaaagcgtcttcgggttctccttcctgctgcctggtgtagatggagacttcctcagt 1020
|||||
Db 1120 CCTCAAGAAAGCGTCTTCCGGTTCCTCTCGTGCTGTGGTAGATGGAGACTTCCTCAGT 1179
QY 1021 gacaccccgagggccctcatcaacgggggagacttccacggcctgcaggtgctggtgggt 1080
|||||
Db 1180 GACACCCGACAGGCCCTCATCAACGCGGGAGACTTCCACGGCCCTGCAGGTGCTGGTGGGT 1239
QY 1081 gtggtgaagatgagggtcgtatttctgttctgttacctggttaccgggccccaggttcagcaaac 1140
|||||
Db 1240 GTGGTGAAGATGAGGGCTCGTATTTCTGTGTTTACGGGGCCCCAGGCTTCAGCAAGAC 1299
QY 1141 aacgagtctcatcagccgggcccaggttcctgcccgggtgcccgtggttccccag 1200
|||||
Db 1300 AACGAGTCTCTCATCAGCCGGCGGAGTCTCTGGCGGGTGCCTGGGGTTCCTCCAG 1359
QY 1201 gtaagtacctggcagccgagggctggtgctcctgattacacagactggctcatcccgag 1260
|||||
Db 1360 GTAAGTACCTGGCAGCCGAGGCTGTGCTCTGATTACACAGACTGGCTGCATCCCGAG 1419
QY 1261 gacccggcacgcctgagggagggcctgagcgatgtggtggcgaccacaaatgctcgtgtgc 1320
|||||
Db 1420 GACCCGGACGCTGTAGGGAGGGCCCTGAGCGATGTGGTGGGGCCACCAATGCTGCTGTC 1479
QY 1321 cccgtggcccgactggtggcgactggctgcccaggtgcccgggttctacgcctacgtc 1380
|||||
Db 1480 CCGGTGGCCGAGCTGGCTGGCGGAGTGGCTGCCAGGGTGCCCGGGTCTACGCCCTACGTC 1539
QY 1381 tttgaacacccgtgtcttcacgcgtctcctgcccctgtggtggtgggtgccccacggctac 1440
|||||
Db 1540 TTTGAACACCGTGTCTTCCACGCTCTCTCTGGCCCTGTGGATGGGGGTGCCCCACGGCTAC 1599
QY 1441 gagatcagttcatctttgggataccccctggaccctctcgaactacacgagagagag 1500
|||||
Db 1600 GAGATCAGTTTCATCTTTGGGATCCCTCTGGACCCCTCTCGAAGTACACGGCAGAGGAG 1659
QY 1501 aaaaatttcgcccagcagactgatggatactgggccaactttgccccgcacaggggatccc 1560
|||||
Db 1660 AAAATCTTGGCCCCAGCGACTGATGCGATACTGGGCCAACTTTGGCCGCACAGGGGATCCC 1719
QY 1561 aatgagccccgagacccccaaaggcccccaaatggcccccgtaacggcgggggtcagcag 1620
|||||
Db 1720 AATGAGCCCCGAGACCCCCAAGGCCCCCAATGGCCCCCGTACACGGCGGGGCTCAGCAG 1779
QY 1621 tacgttagtctggacctgcggccgctggaggtgcggcggggtgcgcgccccagcctgc 1680
|||||
Db 1780 TACGTTAGTCTGGACCTGCGGCCGCTGGAGGTGCGGGGGGTGCGCGGCCCGCCAGGCCTGC 1839
QY 1681 gccttctggaaccgcttccctcccaaatgtcagcgccacc 1722
|||||
Db 1840 GCCTTCTGGAACCGCTTCTCTCCCAAAATGCTCAGCGGCCACC 1881

RESULT 12

US-08-814-095-7

; Sequence 7, Application US/08814095

; Patent No. 6025183

; GENERAL INFORMATION:

;;
;; APPLICANT: Soreq, Hermona
;; APPLICANT: Zakut, Haim
;; APPLICANT: Shani, Moshe
;; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
;; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
;; NUMBER OF SEQUENCES: 7
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: KOHN & ASSOCIATES
;; STREET: 30500 No. 6025183thwestern Highway, Suite 410
;; CITY: Farmington Hills
;; STATE: Michigan
;; COUNTRY: U.S.
;; ZIP: 48334
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/814,095
;; FILING DATE:
;; CLASSIFICATION: 800
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Montgomery, Ilene N.
;; REGISTRATION NUMBER: 38,972
;; REFERENCE/DOCKET NUMBER: 2391.00066
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 539-5050
;; TELEFAX: (248) 539-5055
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 35060 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "Cosmid including ACHE
;; DESCRIPTION: promotor, ACHE gene and ARS gene"
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; POSITION IN GENOME:
;; CHROMOSOME/SEGMENT: 7q22
;; FEATURE:
;; NAME/KEY: promoter
;; LOCATION: 4089..22464
;; OTHER INFORMATION: /function= "ACHE Promotor"
;; OTHER INFORMATION: /standard_name= "ACHE Promotor"
;; FEATURE:
;; NAME/KEY: exon
;; LOCATION: 22465..22537
;; OTHER INFORMATION: /function= "non-translated"
;; OTHER INFORMATION: /gene= "ACHE"
;; OTHER INFORMATION: /number= 1
;; FEATURE:
;; NAME/KEY: exon
;; LOCATION: 24090..25177
;; IDENTIFICATION METHOD: experimental
;; OTHER INFORMATION: /function= "(translation start:
;; OTHER INFORMATION: 24110)"
;; OTHER INFORMATION: /evidence= EXPERIMENTAL
;; OTHER INFORMATION: /gene= "ACHE"
;; OTHER INFORMATION: /number= 2
;; FEATURE:
;; NAME/KEY: exon
;; LOCATION: 25524..26009
;; IDENTIFICATION METHOD: experimental
;; OTHER INFORMATION: /evidence= EXPERIMENTAL
;; OTHER INFORMATION: /gene= "ACHE"
;; OTHER INFORMATION: /number= 3
;; FEATURE:
;; NAME/KEY: exon


```

; MAP POSITION: 3q26
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 160..1881
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /EC_number= 3.1.1.8
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "BCHE"
; OTHER INFORMATION: /note= "butyrylcholinesterase mature peptide"
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 76..159
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: 1..2381
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 76..1884
; US-08-318-826A-9

Query Match 22.0%; Score 380; DB 2; Length 2381;
Best Local Similarity 55.2%; Pred. No. 1.5e-71;
Matches 790; Conservative 0; Mismatches 625; Indels 15; Gaps 2;

QY 146 ggggcattgcctgaagacccccggggccctgtctctgtcttccctggggaaccccccttg 205
Db 200 GAGGGATGAACCTTGACAGTTTTGGTGGCAGCGTAACAGCCCTTCTTGGAAATCCCTATG 259

QY 206 cggagccacccatgggacccccgccttctccaccggagcccaagcagccttggtcag 265
Db 260 CACAGCCACCTCTTGGTAGACTTCGATTCAAAAAGCCACAGTCTCTGACCAAGTGGTCTG 319

QY 266 ggggtgtagacgtacaaaccttccagagtgctgtctgtaccaaatatgtggacacccctatacc 325
Db 320 ATATTGGAAATGCCACANAATATGCAAAATTTCTGTCTGTCAGAACATAGATCAAAAGTTTC 379

QY 326 cagggttttgagggcaccgagatgttggaaccccaaccgtgagctgagcgagcgactgcctgt 385
Db 380 CAGGCTTCCATGGATCAGAGATGTGGAAACCCCAACACTGACCTCAGTGAAGACTGTTTAT 439

QY 386 acctcaacgtgtggacacaccatacccccgccctacatccccccaccctgtcctcgtctgga 445
Db 440 ATCTAAATGTATGGATTCAGACACTAAACCAAAAANAATGCCA---CTGTATTGATATGGA 496

QY 446 tctatgggggtggtctctacagtggggccctcctccttggtgacgtgtacgatgcccgttct 505
Db 497 TTTATGGTGGTGGTTTCAAACTGGAACATCATCTTTTACATGTTTATGATGCAAGTTTC 556

QY 506 tggtagaggccgagaggactgtgctggtgtccatgaactaccgggtgggagcctttggct 565
Db 557 TGGCTCGGGTTGAAAGAGTTATTGTAGTGTCAATGAACATATAGGGTGGGTGCCCTAGGAT 616

QY 566 tcctggccctgcgggggagccgagagagggcccggaatgtggtgtcctcctgagatcagaggc 625
Db 617 TCTTAGCTTTGCCAGGAATCCTGAGGCTCCAGGGAACATGGGTTTATTGTATCAACAGT 676

QY 626 tggccctgcagtggtgcaggagaacgtggagcccttcggtggggtgacccacatcagtgag 685
Db 677 TGGCTCTTCAGTGGGTTCAAAAAAATATAGCAGCCTTTGGTGGAAATCCTAAAAAGTGTA 736

QY 686 cgctgtttgggagagcgcgggagccgcctcgtgggcatgcacctgctgtccccgcca 745
Db 737 CTCTCTTTGGAGAAAGTGCAGGAGCAGCTTCAGTTAGCCTGCATTTGCTTTCTCCTGGAA 796

QY 746 gccggggcctgttccacagggccgtgctgcagagcgggtgcccccaatggaccctgggcca 805
Db 797 GCCATTCAATTGTTACCAGAGCCATTCTGCAAAAGTGGATCCTTTAATGCTCCTTGGGCGG 856

QY 806 cgggtggcatgggagagggcccgctgcagggccacgcagctggcccaacctgtgggctgtc 865
Db 857 TAACATCTCTTTATGAAGCTAGGAACAGACGTTGAACCTTAGCTAAAT----- 905
```

```

QY 866 ctccaggcggcacttggtgggaatgacacagagctgggtagccttccgacacgaccag 925
Db 906 -GACTGGTTGCTCTAGAGAGAATGAGACTGAAATATCAAGTGTCTTAGAAATAAGATC 964

QY 926 cgcaggctcctggtgaaccacgaatggcacgtgctgcctcaagaaagcgtcttccggttct 985
Db 965 CCCAAGAAATTCCTCTGAATGAAGCATTTGTTGTCCTTATGGGACTCCTTTGTCAGTAA 1024

QY 986 ccttcgtgcctgtgtagatggagacttccctcagtgagacacccccagaggccctcatcaacg 1045
Db 1025 ACTTTGGTCCGACCGCTGGATGGTATTCTCTACCTGACATGCCAGACATATTACTTGAAC 1084

QY 1046 cgggagacttccacggcctgcaggtgctggtggtggtggtggtggtggtggtggtggt 1105
Db 1085 TTGGACAATTTAAAAAAACCAGATTTTGGTGGGTGTTAATAAAGATGAAGGGACAGCTT 1144

QY 1106 ttctggtttacggggcccgaggttcacgacaaagacacagagtcctctcatcagccgggccc 1165
Db 1145 TTTTAGTCTATGGTGCCTCTGGCTTCAGCAAGATAACAATAGTATCATATACTAGAAAAAG 1204

QY 1166 agttcctggccgggtgctgggtggttccccaggtaagtacctggcagccgggctg 1225
Db 1205 AATTTCAGGAAGGTTTAAAAATATTTTCCAGGAGTGAGTCAGTTTGGAAAGGAATCCA 1264

QY 1226 tggctcctgcattacacagactggctgcacccccgagaccccgacgcctgagggagggccc 1285
Db 1265 TCCTTTTTCATTACACAGACTGGGTAGATGATCAGAGACCTGAAAAACTACCGTGAGCCT 1324

QY 1286 tgagcagatggtggcgaccacaaatgtcgtgtgccccgtggccccagctggctggcgac 1345
Db 1325 TGGGTGATGTTGTTGGGATATATAATTTTCATATGCCCTGCCTTGGAGTTCCACCAAGAGT 1384

QY 1346 tggctgccaggggtgccgggtgtacgcctacgtctttgaacacccgtgcttccagctct 1405
Db 1385 TCTCAGAATGGGAAATAATGCCCTTTTCTACTATTTTGAACACCGATCCTCCAAACTTC 1444

QY 1406 cctggccccctgtggtgggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1465
Db 1445 CGTGGCCAGAATGGATGGGAGTGATGATGCTGGCTATGAATTTGAATTTGTTGTTTAC 1504

QY 1466 ccctggaccctctcgaaaactacacggcagagggagagaaaaatcttcccccagactgatgc 1525
Db 1505 CTCTGGAAGAAGAGATAATTACACAAAAGCCGAGGAATTTTGAGTAGATCCATAGTGA 1564

QY 1526 gatactgggccaacttggccgcacaggggagatcccaatgagccccagac 1575
Db 1565 AACGGTGGGCAAAATTTGCAAAATATGGGAATCCAAATGAGACTCAGAAC 1614
```

```

RESULT 14
US-08-318-826A-8
; Sequence 8, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckstein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```


; Sequence 1, Application US/093334489

; Patent No. 6291175

; GENERAL INFORMATION:

; APPLICANT: Pierre Sevigny

; APPLICANT: Keith Schappert

; APPLICANT: Heiko Wiesbusch

; TITLE OF INVENTION: METHODS FOR TREATING A NEUROLOGICAL

; TITLE OF INVENTION: DISEASE BY DETERMINING BCHE GENOTYPE

; FILE REFERENCE: 08523/013002

; CURRENT APPLICATION NUMBER: US/09/334,489

; CURRENT FILING DATE: 1999-06-16

; PRIOR APPLICATION NUMBER: 60/089,406

; PRIOR FILING DATE: 1998-06-18

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2416

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-334-489-1

Query Match 22.0%; Score 380; DB 4; Length 2416;
Best Local Similarity 55.2%; Pred. No. 1.5e-71;
Matches 790; Conservative 0; Mismatches 625; Indels 15; Gaps 2;

Qy	146	ggggcattgcctgaagacccccggggccctgtctctgtcttctctgggcacccccctttg	205
Db	254	gaggatgaacttgacagtttttggtggcaggttaacagccctttcttggaattccctatg	313
Qy	206	cggagcaccccatgggacccccgtgcctttctgccaccggagcccaagcagccttggtcag	265
Db	314	cacagccacctcttggtagacttctgattcaaaaagccacacagtccttgaccaagtgtctg	373
Qy	266	gggtgttagacgctacaaccttccagagtgctgtctaccaatatgtggacaccctatacc	325
Db	374	atattggaatgccaaaaatatgcaaatcttctgtctcagacatatagatcaaatgtttc	433
Qy	326	caggttttgagggcaccgagatgtgaaaccccccaaccgtgagctgagcgaggactgcctgt	385
Db	434	caggcttcacatggtacagagatgtggaaccccaaacactgacacctgagtgaaactgttat	493
Qy	386	acctcaacgtgtggacaccataccccggcctacatccccccaccctgtcctcgtctgga	445
Db	494	atctaaatgtatggattccagcacctcaaaccaaaaaatgccca--ctgtattgatatgga	550
Qy	446	tctatgggggtggcttctacagtggggcctcctcctcttgagcgtgtacgatggcgccttct	505
Db	551	tttatgggtgggtgttttcaaaactggaacatcatctttacatgttttatgatggcaagtctc	610
Qy	506	tggtagagggccgagaggactgtgtgtgtgtccatgaactacccgggtgggagcctttggct	565
Db	611	tggctcggttgaaagagttattgtagtgtcaatgaactatagggttggtggcctaggat	670
Qy	566	tcctggccctgccggggagccgagagggccccgggcaatgtgggtctcctggatcagaggc	625
Db	671	tcctagctttgccaggaaatcctgaggtccagggaaactatagggtttatttgatcaacagt	730
Qy	626	tggccctgcagtggtgtcaggagaaactgggcagccttcgggggtgaccccgacatcagtga	685
Db	731	tggctcttcagtggtgtcaaaaaaatatagcagcctttggtggaatcctaaagtgtaa	790
Qy	686	cgtgttttggggagagcgcgggagccctcgttgggtggtggtgacacctgtgtcccccga	745
Db	791	ctctcttttggagaaagtgcaggagcgttctcagtttagcctgcatttgtcttctcctgga	850
Qy	746	gccggggcctgttccacagggccgtgtgtgcagagcgggtgcccccaatggaccctgggcca	805
Db	851	gccattcatgttcaccagagccattctgaaaagtggatccctttaaattgctccttgggcgg	910
Qy	806	cgggtgggcatgggagagggcccgctgcaggggccacgcagctggccacccttggggctgtc	865
Db	911	taacatctcttttatgaagctaggaaacagacgttgaaacttagctaaatt-----	959

Search completed: August 31, 2002, 15:36:07
Job time: 6929 sec

Qy	866	ctccagggcgactggtgggaatgacacagagctggtagcctgccttcggacacgaccag	925
Db	960	-gactggttgcctctagagagaaatgagactgaaataatcaagtgtcttagaaataagatc	1018
Qy	926	cgcaggtcctggtgaaccacgaatggcaactgctgcctcaagaagcgtcttccggttct	985
Db	1019	cccaagaattcttctgaatgaagcatttgtgtccctatgggactccttctgcagtaa	1078
Qy	986	ccttcgtcctgtggtagatggagacttctcactgacacccccagaggccctcatcaacg	1045
Db	1079	actttggtccgaccgtggatggtgatttctcactgacatgccagacatatattctgaac	1138
Qy	1046	cgggagacttccacggcctgcaggtgctggtggtggtggtggtggtggtggtggtggtt	1105
Db	1139	ttggacaatttaaaaaaacccagatttgggtggtggtggtggtggtggtggtggtggtt	1198
Qy	1106	ttctggtttacggggccccaggcttccagcaaaagacaaacagagtcctcactcagccggg	1165
Db	1199	ttttagttctatggtgctcctggcttcagcaaaagataaacaatagataataaactaga	1258
Qy	1166	agttcctggccgggtgctgggtgctgggttccccaggtaagtacctggcagcgaggctg	1225
Db	1259	aatttcaggaaaggtttaaaaaatattttccaggagtgagtgagtttgaaaaggaaatcca	1318
Qy	1226	tggctcctgcattacacagactgggtgcctcccgaggaacccggcagcctgagggagccc	1285
Db	1319	tccttttccattacacagactgggtgagatgatcagagacctgaaaactacccgtgagcc	1378
Qy	1286	tgagcgtatggtggcgacaccacaatgctggtggtggtggtggtggtggtggtggtggt	1345
Db	1379	tgggtgatggtggtgggattataattcattcatatgcctgacctggaggttcaccaaga	1438
Qy	1346	tggctgcccagggtgccccgggtctacgcctacgtcttttgaacacccgtgcttccacg	1405
Db	1439	tctcagaatgggaaataaatgccttttctactatttttgaacacccgatacctccaa	1498
Qy	1406	cctggccctgtggatgggtggtggtggtggtggtggtggtggtggtggtggtggtggt	1465
Db	1499	cgtggccagaatggatgggagtgatggtggtggtggtggtggtggtggtggtggtggt	1558
Qy	1466	ccctggacccctctcgaactacacacgagagggagagaaatcttccgacgactgatgc	1525
Db	1559	ctctggaaagaagagataaattacacaaaagccgaggaattttgagtagatccatagtg	1618
Qy	1526	gatactggccaaactttgcccgcacagggtggtggtggtggtggtggtggtggtggtg	1575
Db	1619	aacggtgggcaaatgttgcaaaaatatgggaatcccaaatgagagactcagaac	1668

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2002, 18:04:30 ; Search time 9324.87 Seconds
(without alignments)
12942.093 Million cell updates/sec

Title: US-09-810-861B-3
Perfect score: 5767
Sequence: 1 agcttgcattgcctgcaggtc.....ctatgaccattgattacgccca 5767

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1:	gb_ba:*
2:	gb_htg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vi:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_vi:*
30:	em_htg_hum:*
31:	em_htg_inv:*
32:	em_htg_other:*
33:	em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Match	Length	ID Description

1	5767	100.0	5767	6	AX275254	AX275254 Sequence
2	2637	45.7	3858	6	AX114873	AX114873 Sequence
3	2637	45.7	4473	12	UCU89964	U89964 Cloning vec
4	2637	45.7	4960	6	AX191663	AX191663 Sequence
5	2637	45.7	5711	6	AX114861	AX114861 Sequence
6	2637	45.7	7332	6	AX191664	AX191664 Sequence
c 7	2637	45.7	8062	6	AX114872	AX114872 Sequence
c 8	2637	45.7	8153	6	AX114871	AX114871 Sequence
c 9	2636	45.7	2746	12	CVPGE4Z	X65305 Cloning vec
c 10	2635.4	45.7	4392	12	CVU47294	U47294 Cloning vec
11	2635.4	45.7	4518	12	XXP35SGFP	U28417 Cloning vec
12	2635.4	45.7	4519	12	U02456	U02456 Cloning vec
13	2635	45.7	4674	12	XXU02437	U02437 Cloning vec
14	2634.4	45.7	5027	12	AF276982	AF276982 Integrati
c 15	2633.8	45.7	5647	1	BACKNB	D29979 Bacillus st
c 16	2633.8	45.7	11612	12	CVU80929	U80929 Cloning vec
c 17	2633.8	45.7	16297	12	CVU75991	U75991 Cloning vec
18	2633.8	45.7	16437	12	AF123770	AF123770 Shuttle v
c 19	2633.8	45.7	18772	12	AF133437	AF133437 Cloning v
c 20	2633.8	45.7	19500	12	CVU75992	U75992 Cloning vec
c 21	2633.4	45.7	2743	12	CVPGE3Z	X65304 Cloning vec
c 22	2632.8	45.7	7387	6	ARI76236	ARI76236 Sequence
c 23	2632.8	45.7	7387	6	AX003207	AX003207 Sequence
c 24	2631.8	45.6	4283	6	I63120	I63120 Sequence 3
c 25	2631.8	45.6	4283	6	I85496	I85496 Sequence 1
c 26	2630.2	45.6	3394	12	AF062083	AF062083 Cloning v
c 27	2630.2	45.6	3394	12	AF083409	AF083409 Cloning v
c 28	2630.2	45.6	3604	12	AF062078	AF062078 Cloning v
c 29	2630.2	45.6	3604	12	AF083407	AF083407 Cloning v
c 30	2630.2	45.6	3617	12	AF062079	AF062079 Cloning v
c 31	2630.2	45.6	3714	12	AF083408	AF083408 Cloning v
c 32	2630.2	45.6	3946	12	AF134573	AF134573 Cloning v
c 33	2630.2	45.6	4026	12	AF062080	AF062080 Cloning v
c 34	2630.2	45.6	4154	12	AF062082	AF062082 Cloning v
c 35	2630.2	45.6	4503	12	AF062081	AF062081 Cloning v
c 36	2630.2	45.6	4795	12	AF134572	AF134572 Cloning v
c 37	2630	45.6	3796	6	E49323	E49323 Infectious
c 38	2630	45.6	5452	12	U02454	U02454 Cloning vec
c 39	2630	45.6	6541	12	XXU02433	U02433 Cloning vec
c 40	2630	45.6	6894	12	XXU02435	U02435 Cloning vec
c 41	2630	45.6	7069	12	U02442	U02442 Cloning vec
c 42	2630	45.6	7164	6	AX195206	AX195206 Sequence
c 43	2630	45.6	7164	12	U02451	U02451 Cloning vec
c 44	2630	45.6	7481	12	XXU02438	U02438 Cloning vec
45	2628.4	45.6	15528	6	A93016	A93016 Sequence 4

ALIGNMENTS

RESULT	1				
AX275254					
LOCUS	AX275254	5767 bp	DNA	linear	PAT 29-OCT-2001
DEFINITION	Sequence 3 from Patent WO0171014.				
ACCESSION	AX275254				
VERSION	AX275254.1	GI:16547674			
KEYWORDS					
SOURCE	synthetic construct.				
ORGANISM	synthetic construct				
	artificial sequence.				
REFERENCE	1 (sites)				
AUTHORS	Mor,T., Soreq,H., Arntzen,C. and Mason,H.				
TITLE	Expression of recombinant human acetylcholinesterase in transgenic plants				
JOURNAL	Patent: WO 0171014-A 3 27-SEP-2001;				
	BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH, INC. (US) ; Mor,				
	Tsafir (US) ; Soreq, Hermona (IL) ; Arntzen, Charles (US) ; Mason,				
	Hugh S. (US)				
FEATURES	Location/Qualifiers				
source	1..5767				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
	/note="plasmid vector pTM034"				

BASE COUNT		1390 a	1521 c	1495 g	1361 t		
ORIGIN							
Query Match		100.0%;		Score 5767;	DB 6;	Length 5767;	
Best Local Similarity		100.0%;		Pred. No. 0;			
Matches 5767;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	agcttgcatgcctgcaggtcaacatggtggagcagcacactctcgtctactccaagaata	60				
Db	1	AGCTTGCAATGCGCTGCAGGTCAACATGGTGGAGCAGCACACTCTCGTCTACTCCAAGAATA	60				
QY	61	tcaaaagatacagtcctcagaagaccagagggtctattgagacttttcaacaaagggtaatat	120				
Db	61	TCAAAGATACAGTCTCAGAAGACCAGAGGGCTATTGAGACTTTTCAACAAAGGGTAATAT	120				
QY	121	cgggaaacctcctcggattccattgccagctatctgtcacttcatacgaaaggacagtag	180				
Db	121	CGGGAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTCATCGAAAGGACAGTAG	180				
QY	181	aaaaggaagatggcttctacaaatgccatcattgcgataaaaggaaggctatcgtttcaag	240				
Db	181	AAAAGGAAGATGGCTTCTACAAATGCCATCATTCGGATAAAGGAAAGGCTATCGTTC AAG	240				
QY	241	aatgcctctaccgacagtggttcccaaagatggacccccacccacaggaacatcgtggaa	300				
Db	241	AATGCCTCTACCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAACATCGTGGAA	300				
QY	301	aaagaagacgttccaaccacgtcttccaagaagtgattgattgattgataaacttttcaaca	360				
Db	301	AAAGAAGACGTTCCAACCACGCTTTCAAAGCAAGTGGATTGATGTGATAAATTTTCAACA	360				
QY	361	aagggtaatatcgggaaacctcctcgtgattccattgccagctatctgtcacttcatcga	420				
Db	361	AAGGGTAATATCGGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTCATCGA	420				
QY	421	aaggacagtgaagaaaggaagatggcttctacaaaatgccatcattgcgataaaggaaaggc	480				
Db	421	AAGGACAGTAGAAAAGGAAGATGGCTTCTACAAATGCCATCATTCGATTAAGGAAAGGC	480				
QY	481	tatcgttccaagaatgcctctaccgacagtggttcccaaagatggacccccacccacagagga	540				
Db	481	TATCGTTCAAGAATGCCTCTACCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGA	540				
QY	541	acatcgtggaagaaagaaagacgttccaaccacgtcttcaaaagcaagtggattgatgtata	600				
Db	541	ACATCGTGGAAAAAGAAAGACGTTCCAACCACGCTTCTCAAAGCAAGTGGATTGATGTGATA	600				
QY	601	tctccactgacgtaaggatgacgcacaatccccactatccttcgcaagaccttccctcta	660				
Db	601	TCTCCACTGACGTAAGGGATGACGCACAATCCCACCTATCCTTCGCAAGACCCCTTCCTCTA	660				
QY	661	tataaggaaqtcatattcatattggagagaccctcgagaattaattctcaacacaacatat	720				
Db	661	TATAAGGAAGTTCATTTTCATTGGAGAGGACCTCGAGAAATTAATTCTCAACACAACATAT	720				
QY	721	acaaacaaacgaatctcaagcaatcaagcatctacttctattgcagcaatttaaatca	780				
Db	721	ACAAAACAACGAATCTCAAGCAATCAAGCATTC TACTTCTATTGCAGCAATTTAAATCA	780				
QY	781	tttcttttaagcaaaagcaattttctgaaaaattttcaccattttacgaacgatagccatg	840				
Db	781	TTTTCTTTTAAAGCAAAAGCAATTTTCTGAAAAATTTTACCATTTTACGAACGATAGCCATG	840				
QY	841	gctcccccgcaagtgtctgctgcacacgccttccctggcttccccactccttctcctcctc	900				
Db	841	GCTCCCCCGCAGTGTCTGCTGCACACGCCCTTCCTGGCTTCCCCACTCCTTCTCCTCCTC	900				
QY	901	ctctggctcctgggtggagagtggtgggctgagggccggaggatgcagagctgctgggtg	960				
Db	901	CTCTGGCTCCTGGGTGGAGGAGTGGGGGCTGAGGGCCGGGAGGATGCAGAGCTGCTGGTG	960				
QY	961	acggtgcgtggtggggtccggctgcggggcattcgcctgaagacccccggggccctgtctct	1020				

Db	961	ACGGTGCCTGGGGCCGGCTGCGGGGCATTGCGCTGAAGACCCCGGGGCCCTGTCTCT	1020				
QY	1021	gctttcctgggcatcccccttgcggagccaccatgggaccccgctcgctttctgccaccg	1080				
Db	1021	GCTTTCCTGGSCATCCCCCTTTGCGGAGCCACCATGGGACCCCGTCGCTTCTGCCACCG	1080				
QY	1081	gagcccaagcagccttggtcaggggtggtagcgcctacaaccttccagagtgtctgtac	1140				
Db	1081	GAGCCCAAGCAGCCTTGGTCAGGGGTGGTAGACGCTACAACCTTCCAGAGTGTCTGTAC	1140				
QY	1141	caatatgtggacacccctataccccagggttttgagggccaccgagatgtgaaaccccaaccgt	1200				
Db	1141	CAATATGTGGACACCCCTATACCCAGGTTTGTAGGGCACCGAGATGTGGAACCCCAACCGT	1200				
QY	1201	gagctgagcaggagactgcctgtacctcaacgtgtggacaccatacccccgccctacatcc	1260				
Db	1201	GAGCTGAGCGAGGACTGCCTGTACTCTAACGTGTGGACACCATAACCCCGGCCTACATCC	1260				
QY	1261	cccacccctgtcctcgtcgtgatctatgggggtggcttctctacagtggggcctcctccttg	1320				
Db	1261	CCCACCCCTGTCTCGTCTGGATCTATGGSGTGGCTTCTACAGTGGGGCTCCTCCTTG	1320				
QY	1321	gacgtgtacgatggccgcttcttgggtacaggccgagaggactgtggtgtccatgaac	1380				
Db	1321	GACGTGTACGATGGCCGCTTCTTGGTACAGSCCGAGAGGACTGTGCTGTGTCCATGAAC	1380				
QY	1381	taccgggtggagccttggcttccctggccctgcggggagccgagagggccccgggcaat	1440				
Db	1381	TACCGGGTGGAGCCTTTGGCTTCTCTGGCCCTGCCGGGAGCCGAGAGGCCCGGGCAAT	1440				
QY	1441	gtgggtctcctggatcagaggctggccctgcagtggtgcaggagaaacgtggcagccttc	1500				
Db	1441	GTGGGTCTCTCGATCAGAGGCTGGCCCTGAGTGGTGCAGGAGAACGTGGCAGCCTTC	1500				
QY	1501	gggggtgaccgcacatcagtgacgctgtttggggagagcgcgggagccctcgtggtgggc	1560				
Db	1501	GGGGGTGACCCGACATCAGTGACGCTGTTTGGGGAGAGCGCGGAGCGCTCGGTGGGC	1560				
QY	1561	atgcacctgtgtccccgccagccgggctgttccacagggccgtgctgcagagcgtt	1620				
Db	1561	ATGCACCTGCTGTCTCCGCCAGCCGGGCGCTGTTCACAGGGCCGTGCTGCAGAGCGGT	1620				
QY	1621	gcccccaatggaccttggccacggtgggcatgggagagggccgtcgcaaggccacgcag	1680				
Db	1621	GCCCCCAATGGACCTTGGCCACGCTGGGATGGGATGGGAGAGGCCCGTCGAGGGCCACGCAG	1680				
QY	1681	ctggcccaaccttgggtgtcctccaggcggtggtggtggtggtggtggtggtggtggtg	1740				
Db	1681	CTGGCCCACTTGTGGGCTGTCTCTCCAGGGCGCACCTGGTGGGAATGACACAGAGCTGTA	1740				
QY	1741	gcctgccttcgggacacgacagcgaggtcctggtgaaacacgaatggcacgtgtgcct	1800				
Db	1741	GCCTGCCTTCGGACACGACCAGCGCAGGTCCTGGTGAACCAACGAATGGCAGCTGCTGCCT	1800				
QY	1801	caagaaagcgtcttcccggttctccttcgtgctgtggtggtggtggtggtggtggtggtg	1860				
Db	1801	CAAGAAAGCGTCTTCCGGTCTCTCCTTCTCGCTGTGGTAGATGGAGACTTCTTCAGTGAC	1860				
QY	1861	acccagagccctcatcaacgcgggagacttccacggcctgcaggtggtggtggtggtggtg	1920				
Db	1861	ACCCAGAGGCCCTCATCAACGCGGGAGACTTCCACGGCCTGCAGGTGCTGGTGGTGTG	1920				
QY	1921	gtgaaggatgagggctcgtattttctggttttacaggggccccaggcttcagcaagaacaac	1980				
Db	1921	GTGAAGGATGAGGGCTCGTATTTTCTGTTTACGGGGCCCCCAGGGCTTCAGCAAGACAAC	1980				
QY	1981	gagtctctcatcagcccggtccgagttcctgcggcggtggtggtggtggtggtggtggtg	2040				
Db	1981	GAGTCTCTCATCAGCCGGGCCGAGTTCTTGGCCGGGTGCGGGTTCGGGGTTCCTCCAGGTA	2040				
QY	2041	agtgacctggcagccgaggtgtggtctctgcattacacagactggctgcacccgaggac	2100				

Db 2041 AGTGACCTGGCAGCGAGGCTGTGGTCTCGCATTTACACAGACTGGCTGCATCCCGAGGAC 2100

Qy 2101 ccggcacgcctgagggagccctgagcgaatggtgggagccacacaatgtcgtgcccc 2160

Db 2101 CCGGCACGCCCTGAGGGAGGCCCTGAGCGATGTGGTGGCGCACCACAATGTCTGTGCCCC 2160

Qy 2161 gtggcccgactggctggcgactggctgccccagggtgcccgggtctacgcctacgtcttt 2220

Db 2161 GTGGCCACGCTGGCTGGCGCACTGGCTGCCACGGGTGCCCCGGGTCTACGCCCTACGTCTTT 2220

Qy 2221 gaacacgctgcttccacgctctcctgcccctgtggatgggggtgccccacggctacgag 2280

Db 2221 GAACACCGTGTCTCCACGCTCTCTGGCCCTGTGGATGGGGGTGCCCCACGGCTACGAG 2280

Qy 2281 atcgagttcatcttbtggatccccctggacccctctgaaacctacacgagagagaaa 2340

Db 2281 ATCGAGTTTCATCTTTGGATCCCCCTGGACCCCTCTGAAACTACACGGCAGAGAGAAA 2340

Qy 2341 atcttcgcccagcgactgacgatactggggccaaacttggccgcacaggggatcccaat 2400

Db 2341 ATCTTGGCCAGCGACTGATCGGATACTGGGCCAACTTTGGCCGACACAGGGGATCCCAAT 2400

Qy 2401 gagccccgagaccccccaagccccacaaatggccccctgtacacggcgggctcagcagtac 2460

Db 2401 GAGCCCCGAGACCCCCAAGSCCCCACAATGGCCCCCTACACGGCGGGGCTCAGCAGTAC 2460

Qy 2461 gttagctctggacctgcccgcctggaaggtgcggggggctgcgcgccacggcctgcgcc 2520

Db 2461 GTTAGTCTGGACCTGCGGCGCTGGAGGTGCGGGGGGGTGGCGCCCGCCAGCCCTGCGCC 2520

Qy 2521 ttctggaaacgccttcctccccaaaattgctcagcgctacctgataggtaccgagctctctc 2580

Db 2521 TTCTGGAAACCGCTTCCTCCCAAAATTGCTCAGCGCTACCTGATAGGTACCGAGCTCTCTC 2580

Qy 2581 aacaaatagctagagtttgctcctctatctatatgtaaaggtatgctgatagcactat 2640

Db 2581 AACAACTAGCTAGAGTTTGCTCCTATCTATATGTAATAAGGTATGCTGATATGCACACTAT 2640

Qy 2641 tcaaataggagcattagctatgttgttaatgtcactttatgttatgtgggtaagtcaacc 2700

Db 2641 TCAAAATAGGAGCATTAGCTATGTTTGTAAATGTCACCTTATGTTATGTGGTAAAGTCACC 2700

Qy 2701 taagacactccacgtacctacgttgtgtctcttaccggctttaataaatcttctgcctc 2760

Db 2701 TAAGACACTCCACGTACCTAGCTGTTGTCTCTTACCGGCTTTAATAAATCTTCTGCCCT 2760

Qy 2761 tgttccatatattactaataattacccttcttcaactaaagaaaattgttatcattaaagtat 2820

Db 2761 TGTTCATATTTACTAATATCCCTTTCTTCACCTAAAGAGAAATTTGTTATCATTAAGTAT 2820

Qy 2821 tagtctttagaacatataggtctttaaattgggtagggttttacaaaattaaactaataaaa 2880

Db 2821 TAGTCTTTAGAACATATAGGTCTTTAAATTGGGTAGGTTTACAAAATTAACATAATAAA 2880

Qy 2881 atgtcaataaaatccacgtggttaaacaataatgcagaaaatcgacgtcgtctattggacega 2940

Db 2881 ATGTCAATAAAATCCACGTGGTTAAACAAATGCAGAAAATCGACGTCGTCTATTGGACCGA 2940

Qy 2941 cagttgctattaaataatggccaccatagtagactgacaaaataaattacctgacaaaca 3000

Db 2941 CAGTTGCTATTAATATATGSSGCCACCATAGTAGACTGACAAATAAATTACCTGACAACA 3000

Qy 3001 tcgtttcactaaataacaaacaaagggagtgcatctttccaggggcatttttgttaata 3060

Db 3001 TCGTTTCACTAAATAACAAACACAAAAGGGAGTGCAATTTTCCAGGGCATTTTGTATAA 3060

Qy 3061 aaaaacagttaaaagggagtgcaatagaaaatataggggtgtggaaaatagtgatttgaca 3120

Db 3061 AAAAACAGTTAAAAGGGAGTGCAATAGAAAATATAGGGGTGTGGAATAGTATTGAGCA 3120

Qy 3121 cgtcttgaagcggaattcactggcgctcgcttttacaaactgctgactgggaaacccctggc 3180

Db 3121 CGTCTTGAAGCGGAATTCACTGGCGCGTCTGTTTACAAACGTCGTGACTGGGAAACCCCTGGC 3180

Qy 3181 gttacccaaacttaatcgccctgcagcacacatccccctttcgccagctggcgtaaatagcgaa 3240

Db 3181 GTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCCGAGCTGGCGTAATAGCGAA 3240

Qy 3241 gagggcccgcacccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgccctg 3300

Db 3241 GAGGCCCCGACCCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGGAATGGCGCTG 3300

Qy 3301 atcggggtatttctccttacgcatctgtgcggtattttcacccgcataatggtgcactctc 3360

Db 3301 ATCGGGTATTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTC 3360

Qy 3361 agtacaaatctgctctgatgcgcgcatagttaagccagccccgcagacaccccgct 3420

Db 3361 AGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCCGCT 3420

Qy 3421 gagcgccccctgacgggcttctgctccccggcatccgccttacagacaagctgtgaccctc 3480

Db 3421 GACGGCCCTGACGGGCTTGTCTGCTCCGGCATCCGCTTACAGACAAGCTGTGACCCGTC 3480

Qy 3481 tcggggagctgcatgtctcagaggttttcaccgtctacccgaaacgcgcgagacgaaag 3540

Db 3481 TCCGGAGCTGCATGTGTCTCAGAGGTTTTCACCGTCTATCACGAAACGCGGAGACGAAAG 3540

Qy 3541 ggctcgtgatacgcctattttataggttaatgcatgataataatggtttcttagacg 3600

Db 3541 GGCCTGCTGATACGCCCTATTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTAGACG 3600

Qy 3601 tcaggtggcaccttttcggggaaaatgtgcgcggaacccctatttggtttatttttctaaata 3660

Db 3601 TCAGGTGGCACCTTTTCGGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAAATA 3660

Qy 3661 cattcaaatatgtatccgcctcatgagacaaataaacctgataaaatgcttcaataatattga 3720

Db 3661 CATTCAAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGA 3720

Qy 3721 aaaaggaagagtatgagtattcaaatctccgtgtgcgccctattcccttttttgcggca 3780

Db 3721 AAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTCGGGCA 3780

Qy 3781 ttttgcccttccctgttttgcaccccaaaaacgcgtggtgaaagttaaagatgctgaagat 3840

Db 3781 TTTTGCCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGAT 3840

Qy 3841 cagttgggtgcacgagtggtttacatcgaactggatctcaacgcggtgaagatcccttgag 3900

Db 3841 CAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAG 3900

Qy 3901 agttttcgccccgaagaacgttttccaatgatgagcacctttttaaagtctgctatgtggc 3960

Db 3901 AGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGC 3960

Qy 3961 gcggtattatcccgctattgacgcgcgggcaagagcaactcggtcgccgcatacacactattct 4020

Db 3961 GCGGTATTATCCCGTATTGACGGCCGGCAAGAGCAACTCGGTCGCGCATACACTATTCT 4020

Qy 4021 cagaatgacttggttgagtactaccagtcacagaaaagcatcttacggatggcatgaca 4080

Db 4021 CAGAAATGACTTGGTTGAGTACTCACAGTCAAGAAAAGCATCTTACGGATGGCATGACA 4080

Qy 4081 gtaagagaattatgcagtgctgccataaccatgagtataacactgcggccaaacttactt 4140

Db 4081 GTAAGAGAATTATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTT 4140

Qy 4141 ctgacaacgatcggaggaccgaaggaggtaaacccgttttttgcacaacatggggatcat 4200

Db 4141 CTGACACGATCGGAGGACCCGAAGSAGCTAAACCGCTTTTTCACACACATGGGGGATCAT 4200

Qy 4201 gtaactcgcccttgatcgttgggaacccggagctgaaatgaagccataaccaaacgacgagcgt 4260

Db 4201 GTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGT 4260

QY 4261 gacacacgagtcgttagcaaatggcaacaacggttcgcgcaaaactatttaactggcgaaacta 4320
Db 4261 GACACCACGATGCCTGTAGCAATGGCAACACGTTGCGCAAACTATTAACTGGCGAACTA 4320
QY 4321 cttactctagcttcccggaacaataatagactggatggagggcgataaaagtgcagga 4380
Db 4321 CTTACTCTAGCTTCCCGGCAACAATTAAATAGACTGGATGGAGGGGATAAAGTTGCAGGA 4380
QY 4381 ccaactctgcgctcgcccttcgggctgcgctggctgggtttattgctgataaaatcctggagccggt 4440
Db 4381 CCACTTCTGGCTCGGCCCTTCCGGCTGGCTGGTATTGCTGATAAATCTGGAGCCGGT 4440
QY 4441 gagcgtgggtctcgcggtatcatctgcagcactggggccagatggtaagccctcccgatatc 4500
Db 4441 GAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATC 4500
QY 4501 gtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgct 4560
Db 4501 GTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCT 4560
QY 4561 gagataggtgcctcactgattaagcatttggttaactgtcagaccaagtcttactcatatata 4620
Db 4561 GAGATAGGTGCTCACTGATTAAAGCATTTGTAAGTGTCACTGTCTAGTGTACTCATATATA 4620
QY 4621 ctttagattgatttaaaacttcatttttaatttaaaggatctagtggaagatcccttttt 4680
Db 4621 CTTTAGATTGATTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTT 4680
QY 4681 gataaatctcatgaccaaataatcccttaacgctgagtttctgctccactgagcgtcagacccc 4740
Db 4681 GATAATCTCATGACCAAATAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCC 4740
QY 4741 gtagaaaaagataaaaggatctcttgagatccctttttctgcgcgtaaatctgctgcttg 4800
Db 4741 GTAGAAAAGATCAAAAGGATCTTCTTGATCCTTTTCTGCGCGTAATCTGCTGCTTG 4800
QY 4801 caaacaaaaaacaccacgctaccagcggtgttctgcgggatacaagagctaccaact 4860
Db 4801 CNAACAAAAAACACCACCGCTACCAGCGGTGTTGTTGCGGGATCAAGAGCTACCAACT 4860
QY 4861 cttttccgaaggttaactggcttcagcagagcgagataccaaataactgtccttctagt 4920
Db 4861 CTTTTCGGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTAGTG 4920
QY 4921 tagccgtagttaggccaccacttcaagaactctgtagcaacgcctacatacctcgctctg 4980
Db 4921 TAGCCGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTG 4980
QY 4981 ctaaatcctgttacacagtggtgctgcagtgccagtgcgataaagtcgtgtcttaccgggtggac 5040
Db 4981 CTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGGATAGTCGTGCTTACCGGTTGGAC 5040
QY 5041 tcaagacgatagtttaccgggataaaggcgagcggtgcgggctgaacggggggttcgtgcaca 5100
Db 5041 TCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGTTTCGTGCACA 5100
QY 5101 cagcccagcttgagcggaacgacctacacccgaactgagataacctacacgctgagctatga 5160
Db 5101 CAGCCCAGCTTGGAGCGGAACGACCTACACCGAACTGAGATACCTTACAGCGTGAGCTATGA 5160
QY 5161 gaaagcgccacgcttcccggaagggaagagcgagaggttatccggttaagcggcagggtc 5220
Db 5161 GAAAGCGCACCGCTTCCCGAAGGGAGAAAGGGCGACAGGTATCCGGTAAGCGGCAGGGTC 5220
QY 5221 ggaacaggagagcgacgagggagcttccagggggaaacgcctcggtatctttatagtcct 5280
Db 5221 GGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACCGCCTGCTATCTTTATAGTCCT 5280
QY 5281 gtcgggttccgacacctctgacttgagcgtcgattttgtgtagctcgtcagggggcg 5340
Db 5281 GTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGG 5340
QY 5341 agcctatggaaaaaacgccagcaacgggcctttttaacgggttcctggccttttgcgtggcct 5400

Db 5341 AGCCTATGGAAAAACGCCAGCAACGCGGCCCTTTTACGGTTCTCTGGCCTTTGCTGGCCT 5400
QY 5401 ttgtctacatgttcttctcctgcgttatccctgattctgtggataacggtattaccgcc 5460
Db 5401 TTTGCTCACATGTTCTTCTCTGCGTTATCCCTCGATTCTGTGGATAACCGTATTACCGCC 5460
QY 5461 tttagtgagctgataccgctcgcgcgacgcgcgaacgacgcgagcgagcgagtcagtgagc 5520
Db 5461 TTTGAGTGAGCTGATACCGCTCGCCGACGCCAACGACCGAGCGCAGGAGTCAGTGAGC 5520
QY 5521 gaggaagcggaagagcgcccaataacgcaaaacgcgcctctcccgcgcttgccgattcat 5580
Db 5521 GAGGAAGCGGAAGAGCGCCCCAATACGCAAAACCGCCTCTCCCCGCGCTTGGCCGATTCA 5580
QY 5581 taatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaatt 5640
Db 5581 TAATGCAGCTGGCAGCAGAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATT 5640
QY 5641 aatgtgagttagctcactcattaggcaccgccaggttttacactttatgcttccggctcgt 5700
Db 5641 AATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTTATGCTTCCGGCTCGT 5700
QY 5701 atgttgtgtggaattgtgagcgataaacaatttcacacacaggaacagctatgacatgat 5760
Db 5701 ATGTTGTGTGAATTGTGAGCGGATAACAAATTTCACACAGGAAACAGCTATGACCATGAT 5760
QY 5761 tacgccca 5767
Db 5761 TACGCCCA 5767

RESULT 2
AX114873
LOCUS AX114873 3858 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 33 from Patent WO0129208.
ACCESSION AX114873
VERSION AX114873.1 GI:14031815
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 3858)
AUTHORS Kuehn,R., von Melchener,H. and Altschmied,J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 33 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)
FEATURES
source
1. 3858
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pRK50"

BASE COUNT 943 a 959 c 994 g 962 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 3858;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcaactggccgtcgtttttacaaactgtgactgggaaaaacccctggcgttaccaccaac 3190
Db 1187 CGAATTCACCTGGCCGTCGTTTTTACAACTGCTGACTGGGAAAACCCCTGGCGTTACCCCAAC 1246
QY 3191 ttaatgccttgacgacacatcccccttttcgccagctggcgtaataagcgaagagggcccgca 3250
Db 1247 TTAATCGCCTTGACGACACATCCCCCTTTCCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 1306
QY 3251 ccgatcgcccttcccaacagttgcgagcctgaatggcgaaatggcgctgatgcggtatt 3310
Db 1307 CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATT 1366
QY 3311 ttctccttacgcatctgtgcgggtatttcacaccgcgcatatggtgcactctcagtacaatct 3370

Qy 5531 aagagcgccaatacgcaaacccctctccccgcgcttgccgattcaatgaatgcagct 5590
|||||
Db 3587 AAGAGCGCCCAATACGCAACCGCCTCTCCCGCGCTTGCGCGATTCAATGAATGCAGCT 3646

Qy 5591 ggcacgacaggttccccgactgaaaggggcagtgagcgcaacgcaaatgaatgtgagtt 5650
|||||
Db 3647 GGCACGACAGGTTTCCCGACTGGAAGCGGCAGTGAGCGCAACGCAATTAATGTGAGTT 3706

Qy 5651 agctcactcattaggcaccgccaggtttacacatttatgettcggctcgatgtttgtgtg 5710
|||||
Db 3707 AGCTCACTCATTAGGCACCCAGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 3766

Qy 5711 gaattgtgagcggaataacaatttcacacaggaacacagctatgaccatgattacgcc 5767
|||||
Db 3767 GAATTGTGAGCGGATAACAATTTACACAGGAACACAGCTATGACCATGATTACGCCA 3823

RESULT 3
UCU89964
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

UCU89964 4473 bp DNA circular SYN 03-SEP-1997
Cloning vector pG5CAT, complete sequence.
U89964
U89964.1 GI:2293130
unidentified cloning vector.
unidentified cloning vector
artificial sequence; vectors.
1 (bases 1 to 4473)
Holtz,A. and Lou,Y.
pG5CAT complete sequence
Unpublished
2 (bases 1 to 4473)
Holtz,A. and Lou,Y.
Direct Submission
Submitted (19-FEB-1997) CLONTECH Laboratories, Inc., 1020 East
Meadow Circle, Palo Alto, CA 94303-4230, USA
This vector can be obtained from CLONTECH Laboratories, Inc., 1020
East Meadow Circle, Palo Alto, CA 94303-4230, USA. To place an
order call (415) 424-8222 or (800) 662-2566, extension 1.
International customers, please contact your local distributor. For
technical information, call (415) 424-8222 or (800) 662-2566,
extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH. If you suspect there is an
error in this sequence, please contact CLONTECH's Technical Support
Department at (415) 424-8222 or (800) 662-2566, extension 3 or
E-mail TECH@CLONTECH.COM.
Location/Qualifiers
1. .4473
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
120. .779
/gene="CAT"
120. .779
/gene="CAT"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="AAB68658.1"
/db_xref="GI:2293131"
/translation="MEKKITGYTTVDISQWHRKEHFEAFQSVAAQCTYNQTVQLDITAF
LKTVKKNKHKEYPAFTHILARLMAHPEFRMAKDGELVIWDSVHPCYTVFHEQTETF
SSLWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGFLENMFVSANPWVSTSFDLNV
ANMNDFFAPVFTMGKYYTQGDVKVLMPLAIQVHHAVCDGFHVGRMLNELQQYCDIEWQGG
A"
2337. .3197
/gene="ampr"
2337. .3197
/gene="ampr"
/function="ampicillin resistance"
/codon_start=1

/transl_table=11
/protein_id="AAB68659.1"
/db_xref="GI:2293132"
/translation="MSIOHFRVALIPFFAAFCPLPFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILESFRPEFPMSTFKVLLCGAVLSRIDAGQEQLGRRIRHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNMGDHVTRL
DRWEPELNEAIPNDERDTPVAMATTLRKLLTGLLELTLASRQQLIDWMEADKVGAPL
LRSLALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSOATMDERNRQIA
EIGASLIKHW"
BASE COUNT 1183 a 1044 c 1046 g 1200 t
ORIGIN

Query Match 45.7%; Score 2637; DB 12; Length 4473;
Best Local Similarity 100.0%; Pred. No.:0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3131 cgaattcactggccgctcggttttacaaacgtcgtagctgggaaaaacccctgacctacccaac 3190
|||||
Db 1735 CGAATTCACTGGCCGCTCGTGTTTTACAACGCTCGTACTGGGAAAACCCCTGGCGTTACCCAAC 1794

Qy 3191 ttaatcgccctgcagcacatccccctttccgacagctggcgtaataagagaagagcccgca 3250
|||||
Db 1795 TTAATCGCCTTGCAGCACATCCCCCTTTCCGCCAGCTGGCGTAATAGCAAGAGGCCCGCA 1854

Qy 3251 ccgatcgcccttcccaacagttgcgcagcctgaatggcgaaatggcgccctgatgcggtatt 3310
|||||
Db 1855 CCGATCGCCCTTCCCAACAGTTGCCGACGCTGAATGGCGAATGGCGCTGATCGCGTATT 1914

Qy 3311 ttctccttacgcatctgtgcgggtattttcacaccgcgcatatggtgcactctcagtacaatct 3370
|||||
Db 1915 TTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCATCTCAGTACAATCT 1974

Qy 3371 gctctgatgcgcgcatagttaagccagcccgacacccgcgaacacccgctgacgcgacct 3430
|||||
Db 1975 GCTCTGATGCCGCATAGTTAAGCCAGCCCGACACCCGCCAACACCCGCTGACGCGCCT 2034

Qy 3431 gacgggcttgtctgctcccgcatccgcttacagacaagctgtgacgctctccgggagct 3490
|||||
Db 2035 GACGGGCTTGTCTGTCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 2094

Qy 3491 gcatgtgcagaggttttcaccgtcatcaccgaaacgcgcgagacgaaaggccctcgtga 3550
|||||
Db 2095 GCATGTGTCAGAGGTTTTCACCGTCAATACCGAAACGCGGAGAGAAAGGGCCTCGTGA 2154

Qy 3551 tacgcctatttttataggttaatgtcatgataataatggtttcttagacgtcaggtggca 3610
|||||
Db 2155 TACGCCCTATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTACAGCTCAGGTGGCA 2214

Qy 3611 cttttcggggaaatgtgcgcgggaacccctatttgtttatttttctaaatacatattcaata 3670
|||||
Db 2215 CTTTTCGGGGAATGTGCGGGAAACCCCTATTGTTTATTTTCTAAATACATTCAAATA 2274

Qy 3671 tgtatccgctcatgagacaaataaaccttgataaaatgcttcaataatattgaaaaaggaaga 3730
|||||
Db 2275 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAGGAAGA 2334

Qy 3731 gtatgattattcaacatttccgctgctgccttattcccttttttgggcattttgccttc 3790
|||||
Db 2335 GTATGAGTATTCAACATTTCCGCTGTCGCCCTTATTCCTTTTTTGGGCATTTTGCCTTC 2394

Qy 3791 ctgtttttgctcaccacagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg 3850
|||||
Db 2395 CTGTTTTTGCTCACCCAGAAACGCTGTGGAAGTAAAAGATGCTGAAGATCAGTTGGGTG 2454

Qy 3851 cacgagtggttacatcgaaactggatctcaacagcggtaagatcccttgagagttttcgcc 3910
|||||
Db 2455 CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC 2514

Qy 3911 ccgaagaacggttttcccaatgatgagcacttttaaagtctctgtatgtgcgcggtattat 3970
|||||
Db 2515 CCGAAGAACGTTTTTCCAATGATGAGCACACTTTTAAAGTCTCTGCTATGTGGCGCGGTATTAT 2574

QY 3971 ccggtattgacgccgggccaagagcaactcggctgcgcgcacatacactattctcagaatgact 4030
 |||||
 Db 2575 CCGGTATTGACCCGGGCAAGAGCAACTCGGTCCGCGCATACACTATTCTCAGAAATGACT 2634
 |||||
 QY 4031 tgggtgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaat 4090
 |||||
 Db 2635 TGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT 2694
 |||||
 QY 4091 tatgcagtgctgccataaccatgagtataaacactgcggccaacttacttctgacaacga 4150
 |||||
 Db 2695 TATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA 2754
 |||||
 QY 4151 tcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaaactcgcc 4210
 |||||
 Db 2755 TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCC 2814
 |||||
 QY 4211 ttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacga 4270
 |||||
 Db 2815 TTGATCGTTGGAAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA 2874
 |||||
 QY 4271 tgcctgtagcaatggcaaacagcttgcgaaactattaaactggcgaaactacttactctag 4330
 |||||
 Db 2875 TGCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 2934
 |||||
 QY 4331 ctcccggaacaattaatagactggatggaggcgggataaagtgcaggaccacttctgc 4390
 |||||
 Db 2935 CTTCCTCGGCAACAATTAAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCACTTCTGC 2994
 |||||
 QY 4391 gctcggcccttccggctggctggtttattgctgataaaactctggagccggatcgtgagcgtgggt 4450
 |||||
 Db 2995 GCTCGGCCCTTCGGCTGGCTGTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT 3054
 |||||
 QY 4451 ctcggggtatcattgcagcactggggccagatggtgaagccctcccgtatcgtagttatct 4510
 |||||
 Db 3055 CTCGGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 3114
 |||||
 QY 4511 acacgacggggagtcaggcaactatggatgaacgaaaatagacagatcgtgagataggtg 4570
 |||||
 Db 3115 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 3174
 |||||
 QY 4571 cctcactgattaagcattggttaactgtcagaccaagtttactcatatatatactttagattg 4630
 |||||
 Db 3175 CCTCACTGATTAAAGCATTTGGTAAGTGTGACAGCAAGTTTACTCATATATATACTTTAGATTG 3234
 |||||
 QY 4631 atttaaaacttcattttttaattaaaggatcttaggtgaagatccctttttgataaactca 4690
 |||||
 Db 3235 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 3294
 |||||
 QY 4691 tgacaaaatacccttaacgtgagttttcgttccactgagcgtcagacccccgtagaaaaa 4750
 |||||
 Db 3295 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAA 3354
 |||||
 QY 4751 tcaaggatcttcttgagatccctttttctgcgcgtaactcgtcgtgtgacgaaacaaaa 4810
 |||||
 Db 3355 TCAAAGGATCTTCTTGAGATCCTTTTCTGCGCGTAACTGCTGCTTGCAAAACAAAAA 3414
 |||||
 QY 4811 aaccaccgtaccagcgggtgtgttgcgggatcaagagctaccaactctttttccga 4870
 |||||
 Db 3415 AACCACCGCTACCAGCGGTGTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGA 3474
 |||||
 QY 4871 aggttaactggcttcagcagagcgagataccaaatactgtccttcttagtgtagccgtagt 4930
 |||||
 Db 3475 AGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGTGAGCCCGTAGT 3534
 |||||
 QY 4931 taggccaccacttcaagaactctgtagcaccgcgcctacatacctcgtctctgctaactcgt 4990
 |||||
 Db 3535 TAGGCCACCACCTTCAAGAACTCTGTAGCACCCGCCCTACATACTCGCTCTGTCTGTAATCCTGT 3594
 |||||
 QY 4991 taccagtgctgctgccagtgcgagataagtcgtgtcttaccgggttggaactcaagacgat 5050
 |||||
 Db 3595 TACCAGTGGCTGCTGCCAGTGGGATAAAGTCGTGTCTTACCGGTTGGACTCAAGACGAT 3654
 |||||
 QY 5051 agttaccggataaaggcgcagcggctcgggctgaacgggggggttcgtgcacacagccccagct 5110
 |||||

Db 3655 AGTTACCGGATAAGGCGCAGCGTCTGGGCTGAACGGGGGTTCTGTGCACACAGCCCAGCT 3714
 |||||
 QY 5111 tggagcgaacgacctacacaccgaactcgagatacctacagcgtgagctatgagaaagcgcca 5170
 |||||
 Db 3715 TGGAGCGAAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 3774
 |||||
 QY 5171 cgttcccgaaaggagaaaaggcgacaggtatccggtgaagcggcagggtcggaacaggag 5230
 |||||
 Db 3775 CGTTCCCGAAGGGAGAAAAGCGGAGGTATCCGGTAAGCGGAGGGTCGGAACAGGAG 3834
 |||||
 QY 5231 agcgacgagggagcttccaggggggaaacgcctgggtatctttatagtcctgtcgggtttc 5290
 |||||
 Db 3835 AGCGCAGGAGGAGCTTCCAGGGGAAACGCCCTGGTATCTTTATAGTCTGTGCGGTTTC 3894
 |||||
 QY 5291 gccacctctgacttgagcgtcgatttttggatgctcgtcaggggggcgagcctatgga 5350
 |||||
 Db 3895 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA 3954
 |||||
 QY 5351 aaaaagccagcaaacgcgccttttttaacgggttcctggccttttgcggccttttgcacaa 5410
 |||||
 Db 3955 AAAACGCCAGCAACGCGGCCCTTTTACGGTTCCTGGCCTTTTGTGGCCTTTTGTCTACA 4014
 |||||
 QY 5411 tgttcttctcgtgttatccctgattctgtggataaacctgataaccgcttttgcagtgag 5470
 |||||
 Db 4015 TGTTCTTCTCTGCGTTATCCCTGATCTGTGGATAACCGTATTACCGCCTTTTGAGTGAG 4074
 |||||
 QY 5471 ctgataccgctcgcgcagccgaacgacccgagcgcagcagtcagtgagcgaagcgg 5530
 |||||
 Db 4075 CTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGAGTCACTAGCGAGGAACGG 4134
 |||||
 QY 5531 aagagcggcccaatacgcacaaaccctctcccgcgcgttggcgaattcattaatgcagct 5590
 |||||
 Db 4135 AAGAGCGCCCAATACGCAAAACCGCCTCTCCCCCGCGTGGCCGATTCAATATGCAGCT 4194
 |||||
 QY 5591 ggcacgacaggtttcccgcactggaagcgggcagtcagtcagcgaactaatgtgagtt 5650
 |||||
 Db 4195 GGCACGACAGGTTTCCCAGACTGGAAAGGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 4254
 |||||
 QY 5651 agctcaactcattaggcacccccaggctttacactttatgcttccggctcgtatgtgtgtg 5710
 |||||
 Db 4255 AGCTCACTCATTAGCACCCCCAGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 4314
 |||||
 QY 5711 gaattgtgagcgggataaacaatttcacaggaacagctatgaccatgattacgcca 5767
 |||||
 Db 4315 GAATTCTGAGCGGATAACAATTTTCACAGAGAAACAGCTATGACCATGATTACGCCA 4371
 |||||

RESULT 4
 AX191663
 LOCUS AX191663 4960 bp DNA linear PAT 15-AUG-2001
 DEFINITION Sequence 29 from Patent WO0149832.
 ACCESSION AX191663
 VERSION AX191663.1 GI:15209844
 KEYWORDS .
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 4960)
 AUTHORS Schwenk,F.
 TITLE Transduction of recombinases for inducible gene targeting
 JOURNAL Patent: WO 0149832-A 29 12-JUL-2001;
 ARTEMIS Pharmaceuticals GmbH (DE)
 FEATURES Location/Qualifiers
 source 1..4960
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="vector pCMV-I-Cre-pA"
 BASE COUNT 1225 a 1213 c 1296 g 1226 t
 ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 4960;

Best Local Similarity 100.0%; Pred. NO. 0;										
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
QY	3131	cgaattcactg	gcgcgtcg	ttttacaag	ctgctgact	gggaaaaac	ccctggc	gttacccaac	3190	
Db	2289	cgaattcactg	gcgcgtcg	ttttacaag	ctgctgact	gggaaaaac	ccctggc	gttacccaac	2348	
QY	3191	ttaatgcctt	gcagcacat	cccccttt	cgccagctg	gcgtaata	agcgaag	agcccgca	3250	
Db	2349	TTAATCGCCT	TGCAGCACAT	CCCCCTTT	CGCCAGCTG	GCCTAATAG	CGAAGAG	CGCCCGCA	2408	
QY	3251	ccgatcgccct	ttcccaacag	ttgcgagc	ctgaatggc	gaatggc	cctgatgc	ggtatt	3310	
Db	2409	CCGATCGCCT	TTCCCAACAG	TTGCGCAGC	CTGAATGGCG	AAATGGCGC	CTGATGCGG	TATT	2468	
QY	3311	ttctccttac	gcgatctgt	gcgggtatt	tcacacgcg	catatggtg	caactctc	agtaacaatct	3370	
Db	2469	TTCTCCTTAC	GCATCTGTG	CGGTATTTCA	CACCCGCAT	ATGGTGC	ACTCTC	AGTACAATCT	2528	
QY	3371	gctctgatgc	gcgcatagt	ttaagccag	cccccgac	ccccgaac	ccccgcgtg	acgcgcct	3430	
Db	2529	GCTCTGATGC	CGCATAGT	TTAGCCAGC	CCCGACAC	CCGCCAAC	CCCGCTG	ACCGCCCT	2588	
QY	3431	gacgggctgt	ctctccggc	atccgcttac	agacaaag	ctgtgac	cgtctc	ccgggagct	3490	
Db	2589	GACGGGCTTG	CTCTCCTCG	CATCCGCTT	ACAGACAAG	CTGTGAC	CGCTCTC	CGGGAGCT	2648	
QY	3491	gcatgtgtc	agaggtttt	ccacgctc	atcacccg	gaacgcgcg	agagaag	ggcctcgtga	3550	
Db	2649	GCATGTGTCA	GAGGTTTTC	CACCGTCAT	CACCGAAAC	CGCGGAG	AGCAAGG	CGCTCGTGA	2708	
QY	3551	tacgcctatt	ttttataggt	taattgcatg	ataataatg	gtttcttag	acgctcag	gtggca	3610	
Db	2709	TACGCCATAT	TTTATAGGTT	AAATGTCATG	ATAATAATG	TTTCTTAG	ACGTCAGG	TGGCA	2768	
QY	3611	cttttcggg	gaaatgtgc	gcggaaac	ccccctatt	gtttattt	ttcttaaa	tacattcaaat	3670	
Db	2769	CTTTTCGGGG	AAATGTGCG	GGAATGTGCG	GAACCCCTAT	TGTTTATTTT	CTAAATAC	ATTCAAATA	2828	
QY	3671	tgtatccgct	catagagaca	ataaacctg	atataac	gtcttcaata	tattgaaaa	aggaaga	3730	
Db	2829	TGTATCCGCT	CATGAGACA	ATAAACCTG	ATATAAATG	CTTCAATA	TATTTGAAAA	AGGAAGA	2888	
QY	3731	gtatgagtatt	ccaacattc	ccgtgtgc	gccttatc	ccccctttt	tgcgccat	tttgccttc	3790	
Db	2889	GTATGATAT	TCAACATTT	CCGTGTG	CGCCCTAT	TCCCTTTT	TGCGGCAT	TTTGCCTTC	2948	
QY	3791	ctgttttgc	tcacccagaa	cgcgtggtg	aaagtaaa	agatgctg	aaagatc	agttgggtg	3850	
Db	2949	CTGTTTTGCT	CACCCAGAA	ACGCTGCTG	AAAGTAA	AAAGATGCT	GAAAGATC	AGTTGGGTG	3008	
QY	3851	cacgagtgggt	tacatcga	actggatc	caacagcg	ggtgaag	atccttg	agagtttgcgc	3910	
Db	3009	CACGAGTGG	TTACATCGA	ACTGGATCT	CAACAGC	GGTAAGATC	CTTGAGAG	TTTTTCGCC	3068	
QY	3911	ccgaagaacg	ttttccaa	tgatgagca	acttttaa	agttctg	ctatg	tggtgcgcgtattat	3970	
Db	3069	CCGAAGAAG	CGTTTCCA	ATGATGAG	CACACTTT	TAAAGTTCT	GTATGTG	GCGCGGTATTAT	3128	
QY	3971	cccgtattg	acgcgggga	agcaactc	ggtgcgcg	catatac	actattct	cagaatgact	4030	
Db	3129	CCCGTATTG	ACGCGCGG	GAAGAGCA	ACTCGGT	CGCCGCAT	ACACTAT	TCTCAGAATGACT	3188	
QY	4031	tggttgagt	actcaccag	tcacagaaa	agcatctt	acggatg	gcgatg	acagtaagagaat	4090	
Db	3189	TGGTTGAGT	ACTCACCA	GTTCACAG	AAAAGCAT	CTTACGG	ATGGSATG	ACAGTAAGAGAAT	3248	
QY	4091	tatgcagtgc	tgccataa	cccatgat	gataacac	ctgcggc	caacttact	cttgacaa	cgca	4150
Db	3249	TATGCAGT	GCTGCCAT	AACCATG	AGTGATA	ACACTG	CGGCCA	ACTTACTTCTG	ACACGA	3308
QY	4151	tcggaggac	cggaagg	agctaacc	gctttttg	cacaacat	gggggatc	atgtaactgc	gcc	4210

Db	3309	TCGGAGGACCGAAGGAGCTAACCGCTTTT	TGCAACAACATGGGGGATCATGTA	ACTCGCC	3368
Qy	4211	ttgatcgttgggaaccggagctgaatgaagccata	ccaaacgacgagcgtgacaccacga	4270	
Db	3369	TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATA	CCAAACGACGAGCGTGACACCACGA	3428	
Qy	4271	tgccgttagcaatggcaacaacgttgcgc	aaactattaaactggcgaaacttactctag	4330	
Db	3429	TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTAT	TAACTGGCGAACTACTTACTCTAG	3488	
Qy	4331	cttccccgcaacaattaatagactggatggagcgga	taaaagt ² tcgaggaccacttctgc	4390	
Db	3489	CTTCCCGCAACAATTAATAGACTGGATGGAGCGGATA	AAAGTTGCAGGACCACCTTCTGC	3548	
Qy	4391	gctcgcccttcgggctggctggtttattgctgata	aaactctggagccggtgagcgtgagcgtg	4450	
Db	3549	GCTCGGCCCTTCGGGCTGGCTGGTTATTGCTGATA	AAATCTGGAGCCCGGTGAGCGTGGGT	3608	
Qy	4451	ctcgcggtatcattgcagcactggggccagatggt	aaagccctcccgta	tcgtagttatct	4510
Db	3609	CTCGCGGTATCATTCAGCACTGGGGCCAGATGGT	TAAGCCCTCCCGTATCGTAGTTATCT	3668	
Qy	4511	acacgacgggagtcaggcaactatggatgaacgaaa	atagacagatcgctgagataggtg	4570	
Db	3669	ACACGACGGGAGTCAGGCAACTATGGATGAACGAA	ATAGACAGATCGCTGAGATAGGTG	3728	
Qy	4571	cctcactgattaaagcattggtaactgtcagacca	agtttactcatatatacttttagattg	4630	
Db	3729	CCTCACTGATTAAGCATTGGTAACCTGTCAGACCA	AGTTTACTCATATATACTTTAGATTG	3788	
Qy	4631	atttaaaacttcattttttaatttaaaaggatctag	gtgaagatcccttttfgataaatctca	4690	
Db	3789	ATTTAAAACTTCATTTTAAATTTAAAGGATCTAG	GTGAAGATCCTTTTTTGATAATCTCA	3848	
Qy	4691	tgacaaaatcccttaacgtgagtttctg	tcca	ctgagcgtcagacccccgtagaaaaga	4750
Db	3849	TGACCAAAATCCCTTAACGTGAGTTTTCGT	TCCACTGAGCGTCAGACCCCGTAGAAAAGA	3908	
Qy	4751	tcaaaggatcttcttgagatccctttttctgcgcg	taatactgctgcttgcaaacacaaaa	4810	
Db	3909	TCAAAGGATCTCTTGAGATCCTTTTTTCTG	CGCGTAATCTGCTGCTTGCAAAACAAAAA	3968	
Qy	4811	aaccaccgctaccagcggtggtttgttgcgggat	caagagctaccaactcttttccga	4870	
Db	3969	AACCACCGCTACAGCGGTGGTTTGTGTCGGGAT	CAAGAGCTACCAACTCTTTTCCGA	4028	
Qy	4871	aggtaactggcttcagcagagcgagata	tacaaaatactgtccttctagtgtagccgtagt	4930	
Db	4029	AGGTAAC	TGGCTTCAGCAGAGCGCAGATACCAAA	TACTGTCTTAGTGAGCCGTAGT	4088
Qy	4931	taggccaccacttcaagaactctgtagcaccgc	cctacatacctcgtcttgctaactcctgt	4990	
Db	4089	TAGGCCACCACCTTCAAGAACTCTGTAGC	ACCGCCTACATACCTCGCTCTGCTAATCCTGT	4148	
Qy	4991	taccagtggctgctgccagtgccagtgcgata	aaagtcgtgtcttaccgggttggaactcaagacgat	5050	
Db	4149	TACCAGTGGCTGCTGCCAGTGGCGATAAGT	CTGTCTTTACCGGGTTGGACTCAAGACGAT	4208	
Qy	5051	agttaccggataaaggcgagcggtcgggtga	acgggggttcgtgcacacagcccagct	5110	
Db	4209	AGTTACCGGATAAAGGCGCAGCGGT	CGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT	4268	
Qy	5111	tgagcgaacgacctacacccgaactgagata	ccctacagcgtgagctatgagaaa	cgcca	5170
Db	4269	TGGAGCGGAACGACCTACACCGAACTGAGATA	CCCTACAGCTGAGCTATGAGAAA	CGCCA	4328
Qy	5171	cgcttcccgaaagggaagggcgacaggtat	ccggtaagcggcgaggtcgggaacaggag	5230	
Db	4329	CGCTTCCCGAAGGGAGAAAGGGCGACAGGT	ATCCGGTAAAGCGGACGAGGTGCGGAACAGGAG	4388	
Qy	5231	agcgcaggggagcttccaggggggaacgc	ctggtatctttatagtagtccgtcggtttc	5290	
Db	4389	AGCGCACGAGGGAGCTTCCAGGGGGAAC	CGCTGGTATCTTTATAGTCTTGTCGGGTTTC	4448	

```
QY 5291 gccacctctgacttgagctcgatgttttltgtgatgctcgtcaggggggagcctatgga 5350
|||||
Db 4449 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA 4508

QY 5351 aaaacgccagcaaacgccccttttacggttcctggccttttgcgtgacctttgtctcaca 5410
|||||
Db 4509 AAAACGCCAGCAACCGCGCCCTTTTACGGTTCCTGGCCTTTTGTGGCCTTTTGTCTACA 4568

QY 5411 tgtttcttctgcgttatccccctgattctgtggataaacctgattaccgcctttgagtga 5470
|||||
Db 4569 TGTTCCTTCCTGCGTTATCCCTCGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG 4628

QY 5471 ctgataccgctcgcgcagcgaacgaccgagcgcagcagtcagtcagtcagtcagtcagtcag 5530
|||||
Db 4629 CTGATACCGCTCGCGCGCAGCGAACGACCGAGCGCAGCGAGTCAGTGACGAGGAAGCGG 4688

QY 5531 aagagcgcccaatacagcaaacgcgcctctccccgcgcgttgccgattcattaatgacgct 5590
|||||
Db 4689 AAGAGCGCCCAATACGCAACCGCCTCTCCCCGCGCTTGGCCGATTCAATTAATGCAGCT 4748

QY 5591 ggcacgacaggtttcccgactggaaagcgggcagtcagtcagtcagtcagtcagtcagtc 5650
|||||
Db 4749 GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 4808

QY 5651 agctcactcattagaccaccaggcctttacacctttatgcttcggcgtcgtatgttgtg 5710
|||||
Db 4809 AGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGCTG 4868

QY 5711 gaattgtgagcggataacaatttcacacaggaacagctatgacctgattacgccaa 5767
|||||
Db 4869 GAATTGTGAGCGGATAACAATTTACACACAGGAACAGCTATGACCATGATTACGCCCA 4925

RESULT 5
AX114861
LOCUS AX114861 5711 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 21 from Patent WO0129208.
ACCESSION AX114861
VERSION AX114861.1 GI:14031803
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 5711)
AUTHORS Kuehn,R., von Melchener,H. and Altschmied,J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 21 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)
FEATURES
source
1..5711
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector PRK65"
BASE COUNT 1329 a 1511 c 1617 g 1254 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 5711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactggccgctcgttttacaaactcgtgactgggaaacccctggcggtaccacaac 3190
|||||
Db 3040 CGAATTCACTGGCCGCTCGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCCAAC 3099

QY 3191 ttaatgccttgacagcacatcccccttccgacgctggcgtaataagcgaagagggcccgca 3250
|||||
Db 3100 TTAATGCGCCTTGACGACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 3159

QY 3251 ccgatcgcccttcccaacagttgcgcagcctgaatggcgaaatggcgctgatgcggtatt 3310
|||||
Db 3160 CCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGNAATGGCGCCTGATGCGGGTATT 3219
```

```
QY 3311 ttctccttacgcacatctgtgcggtattttcacacccgcataatggtgcactctcagtacaatct 3370
|||||
Db 3220 TTCTCCTTACGCATCTGTGCGGTATTTTACACCCGCATATGTTGCACTCTCAGTACAATCT 3279

QY 3371 gctctgatccgcgcatagttaagccagcccccgacaccccgccaacaccccgctgacgcgcct 3430
|||||
Db 3280 GCTCTGATGCCGCATAGTTAAGCAGCCCCCGACACCCGCCCAACACCCGCTGACGCGCCCT 3339

QY 3431 gacgggcttgtctgctccccgcctccgcttacagacaaagctgtgacccgtctcccgggagct 3490
|||||
Db 3340 GACGGGCTTGTCTGCTCCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 3399

QY 3491 gcatgtgtcagaggtttttcacccgtcatcacccgaaacgcgcgagacgaaagggcctcgtga 3550
|||||
Db 3400 GCATGTGTcAGAGGTTTTCACCGTcATcACCGAAACCGCGGAGACGAAAGGCCCTCGTGA 3459

QY 3551 tacgcctatttttataggttaatgtcatgataataatggtttccttagacgtcagtggtga 3610
|||||
Db 3460 TACGCGCTATTTTATAGGTTAATGTcATGATAATAATGGTTTCTTAGACGTcAGGTGGCA 3519

QY 3611 ctttcggggaaaatgtgcgcggaacccctatttgtttatttttttaataacattcacaata 3670
|||||
Db 3520 CTTTTCGGGGAAAATGTGCGGGAAACCCCTATTtTGTtTATTtTCTTAATtACATTCAAATA 3579

QY 3671 tgtatccgctcatgagacaaataaccctgataaaatgcttcaataataattgaaaaaggaaga 3730
|||||
Db 3580 TGtATCCGCTCATGAGACAATAAACCTGATAAAATGCTTCAATAATATtTGAAAAAGGAAGA 3639

QY 3731 gtatgagtattcaacatttcctgctgcgcccttattcccttttttgcggcattttgccttc 3790
|||||
Db 3640 GTATGAGTATTCAACATTTCCGtGTGCGCCCTATTtCCCTTTTTTGGGGCATTTTGGCCTTC 3699

QY 3791 ctgttttctgcacccagaaaacgctgggtgaaaagtataaagtgcgtgaagatcagttgggtg 3850
|||||
Db 3700 CTGTtTTTGCTcACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATcAGTTGGGTG 3759

QY 3851 cacgagtgggttacatcgaactggatctcaacagcggtaagatcccttgagagttttcgcc 3910
|||||
Db 3760 CACGAGTGGGTtTACATCGAACTGGATCTCAACAGCGGTAAGATCTTtGAGAGTtTTTCGCC 3819

QY 3911 ccgaagaacgttttcccaatgatgagcaccttttaaagttctgtatgtggcggtattat 3970
|||||
Db 3820 CCGAAGAACGTTtTCCAATGATGAGCACTTTtTAAAGTTCTGCTATGTGGCGGGTATTAT 3879

QY 3971 ccgatttgacgcgggcaagagcaactcgtgcgcgcatacacactattctcagaatgact 4030
|||||
Db 3880 CCCGATTtGACGCGCGGCAAGAGCAACTCGGTGCGCCGCATACACTATTtTCTCAGAATGACT 3939

QY 4031 tggttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaat 4090
|||||
Db 3940 TGGTTGAGTACTcACCAGTcACAGAAAAGCATCTTtACGGATGGCATGACAGTAAGAGAAAT 3999

QY 4091 tatgcagtgtgccataaaccatgagtgataaacactgcggccaacttacttctgacaaacga 4150
|||||
Db 4000 TATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGGCCAACCTTACTTCTGACAACGA 4059

QY 4151 tcggaggaccggaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgcc 4210
|||||
Db 4060 TCGGAGGACCGGAAGGAGCTAACCGCTTTtTtTGCAACAACATGGGGATCATGTAACTCGCC 4119

QY 4211 ttgatcgttgggaacccggagctgaatgaagccataccaaaacgacgagcgtgacacccacga 4270
|||||
Db 4120 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCCAACGACGAGCGGTGACACCACGA 4179

QY 4271 tgcctgtagcaaatggcaacaacggttgcgcaaaactattaaactggcgaaacttacttactag 4330
|||||
Db 4180 TGCTGTAGCAATGGCAACACGTTGGCGAAACTATTtAACTGGCGAACTACTTACTCTAG 4239

QY 4331 cttcccgccaacaattaatagactggatggagggcggaataaagttgcaggaccacttctgc 4390
|||||
Db 4240 CTTCCCGGCAACAATTAATAGACTGGATGGAGGGCGGATAAAGTTGCAGGACCACCTTCTGC 4299
```


Qy 4391 gctcgcccttccggctggtgttattgctgataaaatctgagccggtgagcgtgggt 4450
|||||
Db 4300 GCTCGGCCCTTCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT 4359
|||||
Qy 4451 ctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagtattct 4510
|||||
Db 4360 CTCGGGTATCATTCAGCACACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 4419
|||||
Qy 4511 acacgcggggagtcaggcaactatggatgaacgaaatagacagatcgctgagatagtg 4570
|||||
Db 4420 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 4479
|||||
Qy 4571 cctcactgattaaagcattggttaactgtcagaccaagtcttactatatactatttagattg 4630
|||||
Db 4480 CCTCACTGATTAAAGCATTGGTAACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTG 4539
|||||
Qy 4631 atttaaaacttcattttttaatttaaaaggatctaggtgaagatcctttttgataatctca 4690
|||||
Db 4540 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCA 4599
|||||
Qy 4691 tgaccbaaaatcccttaacgtgagtttctgcttccactgagcgtcagaccccgtagaaaaaga 4750
|||||
Db 4600 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA 4659
|||||
Qy 4751 tcaaaagatctcttgagatccctttttctgcgcgtaactgtcgtgcttgcaaacaaaaa 4810
|||||
Db 4660 TCAAAGGATCTCTTGAGATCCCTTTTCTGCGGTAATCTGCTGCTTGCAAAACAAAA 4719
|||||
Qy 4811 aaccacgctaccagcgggtgtgtttgttcgggatcaagagctaccacactcttttccga 4870
|||||
Db 4720 AACCACCGCTACCAGCGGTGTTGTGTTGCGGATCAAGAGCTACCAACTCTTTTCCGA 4779
|||||
Qy 4871 aggttaactggcttcagcagagcgcagatacacaactactgtccttctagttagccgtagt 4930
|||||
Db 4780 AGGTAACCTGGCTTCAGCAGAGCGCATACCAATACTGTCTCTTAGTGTAGCCGTAGT 4839
|||||
Qy 4931 tagggccaccactcaagaactctgtagcaccgcctacatacctcgtctgtaactcctgt 4990
|||||
Db 4840 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT 4899
|||||
Qy 4991 taccagtggctgctgccagtgccgataaagtcgtgtcttaccgggttggaactcaagacgat 5050
|||||
Db 4900 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACC GG GTGGACTCAAGACGAT 4959
|||||
Qy 5051 agttacgggataaaggcgcagcgggtcgggctgaacgggggttcgtgcacacagcccagct 5110
|||||
Db 4960 AGTTACCGGATAAAGCGCAGCGGTGGGCTGAACGGGGGTTTCGTGCACACAGCCCCAGCT 5019
|||||
Qy 5111 tggagcgaacgacctacacccgaactgagatacctacagcgtgagctatgagaaagcgcca 5170
|||||
Db 5020 TGGAGCGAAGACGCTACACCGAAGTTCAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 5079
|||||
Qy 5171 cgcttccgaaggagaaaggcggacaggtatccggttaagcggcagggtcggaacaggag 5230
|||||
Db 5080 CGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGACAGGGTCGGAACAGGAG 5139
|||||
Qy 5231 agcgacgaggaggagcttccaggggaaacgcctgggtatctttatagtcctgctgggtttc 5290
|||||
Db 5140 AGCGCACGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTTATAGTCCTGTGCGGGTTTC 5199
|||||
Qy 5291 gccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatgga 5350
|||||
Db 5200 GCCACCTCTGACTTGAGCGTCGATTTTGTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGA 5259
|||||
Qy 5351 aaaacgcagcaacgcggcctttttaagggttcctggccttttgccttttgcctcaca 5410
|||||
Db 5260 AAAACGCCAGCAACCGGCCCTTTTACGGTTCTTGCCCTTTTGTGGCCCTTTTGTGTCACA 5319
|||||
Qy 5411 tgttcttctgctgttatccctgattctgtggataaccgtattaccgcctttgagtgag 5470
|||||
Db 5320 TGTTCCTTCCTGCGTTATCCCTGATTCTGTGGATAACCGGTATTACCGCCTTTTGAGTGAG 5379
|||||
Qy 5471 ctgatacagctcgccgcagccgaacgacccgagcgcagcgagtcagtgagcgaagcgg 5530
|||||

Db 5380 CTGATACCGCTCGCCGACGCGGAACGACGAGCGGAGCTAGTGAGCGAGGAAGCGG 5439
|||||
Qy 5531 aagagcgcacaatacgaaaaaccgcctctccccgcgcttgccgattcattaatgcagct 5590
|||||
Db 5440 AAGAGCGCCCAATACGAAACCGCCTCTCCCGCGCTTGCCGATTCAATTAATGCAGCT 5499
|||||
Qy 5591 ggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagtt 5650
|||||
Db 5500 GGCACGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 5559
|||||
Qy 5651 agctcactcattaggcaccgccaggcttttacactttatgcttccggctcgtatgtgtg 5710
|||||
Db 5560 AGCTCACTCATTAGGCACCCCGAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTG 5619
|||||
Qy 5711 gaattgtgagcgggataaacaattttcacacaggaacagctatgacctgattacgcca 5767
|||||
Db 5620 GAATTGTAGCGGGATAACAATTTTACACAGGAAACAGCTATGACCATGATTACGCCA 5676
|||||
RESULT 6
AX191664
LOCUS AX191664 7332 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 30 from Patent WO0149832.
ACCESSION AX191664
VERSION AX191664.1 GI:15209845
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 7332)
AUTHORS Schwenk,F.
TITLE Transduction of recombinases for inducible gene targeting
JOURNAL Patent: WO 0149832-A 30 12-JUL-2001;
ARTEMIS Pharmaceuticals GmbH (DE)
FEATURES
Location/Qualifiers
source 1..7332
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pCMV-I-beta-pA"
BASE COUNT 1718 a 1895 c 1964 g 1755 t
ORIGIN
Query Match 45.7%; Score 2637; DB 6; Length 7332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3131 cgaattcactggccgctcgttttacacgctgtagctgggaaaaccctggcttaccccaac 3190
|||||
Db 4661 CGAATTCACTGGCCGTCGTTTACAACTGCTGACTGGGAAAACCCCTGGCGTTACCCAAC 4720
|||||
Qy 3191 ttaatcgcttgacgacacatcccccttttcgcccagctggcgtaataagcgaaggcccgca 3250
|||||
Db 4721 TTAATCGCCTTGACGACACATCCCCCTTTCGCCAGCTGGCGTAATGGCGCAATGGCGCTGATGCGGTATT 4780
|||||
Qy 3251 ccgatcgcccttccccaaacagttgcgagccctgaatggcgcaatggcgctgatgcggtatt 3310
|||||
Db 4781 CCGATCGCCCTTCCCAACAGTTGCGCAGCCCTGAATGGCGCAATGGCGCTGATGCGGTATT 4840
|||||
Qy 3311 ttctccttacgcatctgtgcggtattttcacaccgcataatggtgcactctcagtacaatct 3370
|||||
Db 4841 TTCTCCTTACGCATCTGTGCGGTATTTTACACCCGCATATGGTGCACCTCTCAGTACAATCT 4900
|||||
Qy 3371 gctctgatccgcatagttaagccagcccgacaccccgcaacacccgctgacgcgccct 3430
|||||
Db 4901 GCTCTGATGCGGCATAGTTAAGCCAGCCCGGACACCCCGCAACACCCGCTGACGCGCCCT 4960
|||||
Qy 3431 gacgggctgtctgtccccggcatcccgcttacagacaagctgtgacgctctccgggagct 3490
|||||
Db 4961 GACGGGCTGTCTGTCTCCCGGCATCCCGTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 5020
|||||
Qy 3491 gcattgtctcagagggttttccacgctcatcaccgaaacgcgcgagacgaaaggcctcgtga 3550
|||||

Db	5021	GCATGTGT	CAGAGGTTTTCACCGT	CATCACCGAAACGGCGAGACGAAGGCCCTCGTGA	5081
QY	3551	tacgcctattttatagg	ttaatg	tcatgataataatggttttcttagacgtcaggtggca	3610
Db	5081	TACGCCATATTTTATAGGTTAATGTCATGATAAATAATGGTTTCTTAGACGT	CAGGTGGCA	5140	
QY	3611	cttttcggggaaatgtgcgcgggaacccctatttggttatttttcttaaaatacattcaaat	3670		
Db	5141	CTTTTCGGGGAAATGTGCGCGGAACCCCTATTGTATTATTTTCTAAATACATTCAAATA	5200		
QY	3671	tgtatccgctcatgagacaaataaacccctgataaaatgcttcaataatatgaaaaagggaaga	3730		
Db	5201	TGTATCCGCTCATGAGACAAATAACCCCTGATAAAATGCTTCAATAATAATTGAAAAAAGGAAGA	5260		
QY	3731	gtatgagtattcaacattccggtgcgccttattcccttttttgcggcaatttgccttc	3790		
Db	5261	GTATGAGTATTCAACATTTCCGTGTGCCCCATTATTCCTTTTTCGGGCATTTTGCCTTTC	5320		
QY	3791	ctggttttgcctcccgagaaacgcgtggtgaaagtaaaagatgctgaagatcagtcagttgggtg	3850		
Db	5321	CTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG	5380		
QY	3851	cacgagtggtttacatcgcgaactggatctcaacagcggtaagatccttgagagtttttcgcc	3910		
Db	5381	CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATTCCTTGAGAGTTTTCGCC	5440		
QY	3911	ccgaagaacgttttccaaatgatgagcaccttttaaagttctgctatgtggcgcggtattat	3970		
Db	5441	CCGAAGAACGTTTCCAAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGGTATTAT	5500		
QY	3971	cccgtattgacgcggggcaagagcaactcggtgcgcgcatacacataattctcagaatgact	4030		
Db	5501	CCCGTATTGACGCGGGCAAGAGCACTCGGTGCGCGCATACACTATTCTCAGAATGACT	5560		
QY	4031	tggttgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagaaat	4090		
Db	5561	TGGTTGAGTACTACCAGTCAAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT	5620		
QY	4091	tatgcagtctgccataaccatgagtataaacactgcggccaaacttacttctgacaaacga	4150		
Db	5621	TATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAACGA	5680		
QY	4151	tcggaggaccgaaggagctaaccgcttttttgcaacaacatgggggatcatgtaaactcgcc	4210		
Db	5681	TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCC	5740		
QY	4211	ttgatcgttgggaacccggagctgaatgaagccataccaaacgacgagcgtgacaccacga	4270		
Db	5741	TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA	5800		
QY	4271	tgccgttagcaatggcaacaacgcttgcgcgaactattaaactggcgaactacttactctag	4330		
Db	5801	TGCCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG	5860		
QY	4331	cttcccggaacaataatagactggatggaggcgggataaaagtgcaggacacactctgc	4390		
Db	5861	CTTCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAAGTTGCAGGACCACTTCTGC	5920		
QY	4391	gctcggcccttcggcgtggctgggtttatttgctgataaaatctggagccggtgagcgtgggt	4450		
Db	5921	GCTCGGCCCTTCCGSGTGGCTGGTTTATTGCTGATAAATCTGGAGCCGCTGACGCTGGGT	5980		
QY	4451	ctcgcgggtatcatctgcagcactggggccagatggtaagccctcccgtatcgtagttatct	4510		
Db	5981	CTCGCGGTATCATTCGACGACTGGGGCCAGATGGTAAGCCCCCTCCCGTATCTGTTATCT	6040		
QY	4511	acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg	4570		
Db	6041	ACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG	6100		
QY	4571	cctcactgattaaagcattggtaactgtcagaccgaagtttactcatatatactttagattg	4630		

Db	6101	CTCACTGATTAAGCATTTGGTAACCTGTGACACCAAGTTTACTCATATATACTTTAGATTG	6160
QY	4631	atttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttttgataatctca	4690
Db	6161	ATTTAANAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA	6220
QY	4691	tgaccaaaaatcccttaacgtgagtttttcgttccactgagctcagacccccgtagaaaaa	4750
Db	6221	TGACCAAAATCCCTTTAACGTGAGTTTTCGTTCCACTGACGCTCAGACCCCGTAGAAAAGA	6280
QY	4751	tcaaaaggatcttcttgagatccctttttttctgcgcgtaaatctgctgctgctgcaaaacaaaa	4810
Db	6281	TCAAAGGATCTTCTTGAGATCCTTTTTCGCGCGTAATCTGCTGCTTGCAAAACA AAAA	6340
QY	4811	aaccacgcgctaccagcggtgtgtgtgttgcgggatcaagagctaccaactctttttccga	4870
Db	6341	AACCACCGCTACCAGCGGTGGTTTGTGTTGCCGATCAAGAGCTACCAACTCTTTTTCGGA	6400
QY	4871	aggttaactggcttcagcagagcgagataccaaaatactgtccttctagtgtagccgtagt	4930
Db	6401	AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCCGTAGT	6460
QY	4931	taggcaccacttcaagaaactctgtagcaccgcctacatacctcgctctgctaactcctgt	4990
Db	6461	TAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCCTGT	6520
QY	4991	taccagtggctgctgccagtgccgataaagtcgtgctctaccgggttggaactcaagacgat	5050
Db	6521	TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGAT	6580
QY	5051	agttaccgggataaaggcgagcggtcggtgaacgggggttcgtgcacacagccccagct	5110
Db	6581	AGTTACCGGATAAGGCGCAGCGTGGGGCTGAACGCGGGGGTTTCGTGCACACAGCCCCAGCT	6640
QY	5111	tggagcgaacgacctacacggaactgagatacctacagcgtgagctatgagaaaagcgcca	5170
Db	6641	TGGAGCGAAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA	6700
QY	5171	cgcttcggaaggagaaaaggcgacaggtatccggtaagcggcagggtcggaaacaggag	5230
Db	6701	CGCTTCCGAAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAG	6760
QY	5231	agcgcaaggaggagctccagggggaaacgccttggtatctttatagtcctgtcgggttcc	5290
Db	6761	AGCGCACGAGGGAGCTTCCAGGGGGAACCGCTGTATCTTTATAGTCTGTCTCGGGTTTC	6820
QY	5291	gccacctctgacttgagcgtcgatttttgtgagtcgctgcagggggcgagccctatgga	5350
Db	6821	GCCACCTCTGACTTGACGCTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCCTATGGA	6880
QY	5351	aaaaagccagcaacgcggccttttttacgggttccttgcccttttgcggccttttgcctaca	5410
Db	6881	AAAACGCCAGCAACGCGGCCTTTTACGGTTCCTGGCCTTTTGTGCTGGCCTTTTGTCTACA	6940
QY	5411	tgttcttctcgcgttatccctgattctgtggataaccgtattaccgcctttgagtgag	5470
Db	6941	TGTTCTTCTCTGCGTTATCCCCTGATTCTGTGGATAACCGGTATTACCGCCTTTTGTGAGTGAG	7000
QY	5471	ctgatacgcctcgccgcagccgaacgacgcagcgagtcagtgagcgagggaagcgg	5530
Db	7001	CTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGG	7060
QY	5531	aagagcgccaatacgcgaacgcgcctctcccgcggttgccgattcattaaatgcagct	5590
Db	7061	AAGAGCGCCCAATACGCAACCGCCTCTCCCGCGCGTTGGCCGATTTCATTAATGCAGCT	7120
QY	5591	ggcacgacagggtttcccgactggaaagcgggcgagtgagcgcaacgcaattaatgtgagtt	5650
Db	7121	GGCACGACAGGTTTCCCGACTGGAAAGCGGGCGAGTGAGCGCAACGCAATTAATGTGAGTT	7180
QY	5651	agctcactcaattaggcaccaggcctttacactttatgcttccggctcgtatgttgtg	5710
Db	7181	AGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG	7240

QY 5711 gaattgtgagcggataacaatttcacacaggaagaaacagctatgaccatgattacgccca 5767
|||||
Db 7241 GAATTGTGAGCGGATACAAATTTACACACAGGAAACAGCTATGACCATGATTACGCCA 7297

RESULT 7
AX114872/c
LOCUS AX114872 8062 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 32 from Patent WO0129208.
ACCESSION AX114872
VERSION AX114872.1 GI:14031814
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 8062)
AUTHORS Kuehn,R., von Melchener,H. and Altschmied,J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 32 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)
FEATURES
source
1. .8062
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector pRK76"
BASE COUNT 1960 a 2036 c 2121 g 1945 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 8062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactggcgcgtgttttacaacgtcgtagctgggaaacccctggcttacccaac 3190
|||||
Db 8045 CGAATTCACTGGCGTCGTTTACAACGTCGTGACTGGGAAACCCCTGGCGTTACCCAAC 7986

QY 3191 ttaatgccttgagcacacatccccctttcgccagctggcgtaatagcgaagagcccgca 3250
|||||
Db 7985 TTAATCGCCTTGAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 7926

QY 3251 ccgatacgccttcccaacagttgcgcagcctgaatggcgaaatggcgccctgatgcggtatt 3310
|||||
Db 7925 CCGATCGCCCTTCCCAACAGTTGGCGAGCCTGAATGGCGAATGGCGCCTGATGCGGTATT 7866

QY 3311 ttctccttaacgcatctgtgcggtatttcaacccgcataatggtgcactctcagtaaatct 3370
|||||
Db 7865 TTCTCCTTACGCATCTGTGCGGTATTTTCACCCGCATATGGTGCACCTCTCAGTACAATCT 7806

QY 3371 gctctgatccgcatagtttaagccagcccgacacccgccaaaccccgctgacgcgcct 3430
|||||
Db 7805 GCTCTGATGCCGATAGTTAAGCCAGCCCGACACCCGCCAACACCCCGCTGACGCGCCT 7746

QY 3431 gacgggcttgtctgctcccgcatccgcttacagacaagctgtgacgctctccggagct 3490
|||||
Db 7745 GACGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT 7686

QY 3491 gcatgtgtcagaggttttcacccgtcatcaccgaaacgcgcgagacgaaaggccctcgtga 3550
|||||
Db 7685 GCATGTGTCAGAGGTTTTCACCGTTCATACCGAAACGCCGGAGACGAAAGGGCCCTCGTGA 7626

QY 3551 tacgcctatttttatagggttaatgtcatgataataatgggttctttagacgctcaggtggca 3610
|||||
Db 7625 TACGCCTATTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTAGACGTCAGGTGGCA 7566

QY 3611 cttttcggggaatgtgcgcggaacccctatttgtttattttcttaaatacattcaata 3670
|||||
Db 7565 CTTTTCGGGGAATGTGCGGGAACCCCTATTGTTTATTTTCTTAATACATTCAAATA 7506

QY 3671 tgtatccgctcatgagacaataaaccttgataaaatgcttcaataatattgaaaaaggaaga 3730
|||||
Db 7505 TGTATCCGCTCATGAGACAATAAACCCCTGATAAATGCTTCAATAATAATTGAAAAAGGAAGA 7446

QY 3731 gtatgagtattcaacatttccgtgtcgcccttattcccttttttgcgcgaattttgccttc 3790
|||||
Db 7445 GTATGAGTATTCAACATTTCCGTGTGCGCCCTATTCCCTTTTTCGGCATTTTCCTTC 7386

QY 3791 ctgttttgcaccagaaaacgcgtggtgaagtaaaaagatgctgtaagatcagttgggtg 3850
|||||
Db 7385 CTGTTTGTGCTCACCAGAAACGCTGGTGAAGTAAAAGATGCTGAAGATCAGTTGGGTG 7326

QY 3851 cacgagtgggttacatcgaaactggatctcaacagcggtaagatcccttgagagttttcgc 3910
|||||
Db 7325 CACGAGTGGGTTACATCGAACTGGATCTCAACACGGGTAAAGATCCTTGAGAGTTTTCGCC 7266

QY 3911 ccgaagaacgttttccaatgatgagcaccttttaaaagtctctgtatgtggtgcggtattat 3970
|||||
Db 7265 CCGAAGAACGTTTCCAATGATGAGCACCTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 7206

QY 3971 cccgtattgacgcgggcaagagcaactcggtcgccgcatacacactatttcagaatgact 4030
|||||
Db 7205 CCCGTATTGACGCGGGCAAGAGCAACTCGTGCGCCGCATACACTATTCTCAGAATGACT 7146

QY 4031 tggttgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaat 4090
|||||
Db 7145 TGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT 7086

QY 4091 tatgcagtgtgccataaaccatgagtataaacactgcggccaacttacttcttgacaacga 4150
|||||
Db 7085 TATGCAGTGCTGCCATAACCATGAGTGATGATAACACTGCGGCCAACTTACTTCTGACAACGA 7026

QY 4151 tcggaggaccgaaggagtaaacgcgttttttgcaacaacatgggggatacatgttaactcgcc 4210
|||||
Db 7025 TCGGAGGACCGAAGGAGCTAACCCGCTTTTTTTCACAACAACATGGGGGATCATGTAACCTCGCC 6966

QY 4211 ttgatcgttgggaaccggagctgaatgaagccaatacacaacgacgagcgtgacaccacga 4270
|||||
Db 6965 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGA 6906

QY 4271 tgcctgtagcaatggcaacaacgcttgcgcaaaactattaaactggcgaaactacttactctag 4330
|||||
Db 6905 TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 6846

QY 4331 cttcccggaacaataatagactggatggaggcgggataaaaagtgcaggaccacttctgc 4390
|||||
Db 6845 CTTCCCGGCAACAATTAATAGACTGGATGGAGCGGGATAAAAGTTGCAGGACCACCTTCTGC 6786

QY 4391 gctcggcccttcggcgtggctggtttattgtctgataaaatctggagccggtgagcgtgggt 4450
|||||
Db 6785 GCTCGGCCCTTTCGGCTGGCTGGTTTATTGTGTGATAAATCTGGAGCCCGTGAGCGTGGGT 6726

QY 4451 ctcgcgggtatcattgcagcactggggccagatgggtaagccctccgtatcgttagttatct 4510
|||||
Db 6725 CTCGGGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6666

QY 4511 acacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg 4570
|||||
Db 6665 ACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 6606

QY 4571 cctcactgattaagcattggtaactgtcagaccaagtcttactcatatatactttagattg 4630
|||||
Db 6605 CCTCACTGATTAAAGCATTTGTAACGTGTACACCAAGTTTACTCATATATACTTTTAGATTG 6546

QY 4631 atttaaaacttcatttttaatttaaaggatctaggtgaagatcccttttgataatctca 4690
|||||
Db 6545 ATTTAAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCA 6486

QY 4691 tgacaaaaatcccttaacgtgagtttttcctccactgagcgtcagaccccgtagaaaaaga 4750
|||||
Db 6485 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCTCCACTGAGCGCTCAGACCCCGTAGAAAAGA 6426

QY 4751 tcaaaggatcttcttgagatcccttttttcttgcgcgtaaatctgctgcttgcatacaaaaa 4810
|||||
Db 6425 TCAAAGGATCTTCTTGAGATCCCTTTTTTTTCTGCGCGTAATCTGCTGCTTGCATAACAAAAA 6366

QY 4811 aaccaccgctaccagcggtggtttgttgcgggataaagagctaccaactctttttccga 4870
|||||
Db 6365 AACCAACCGCTACCAAGCGGTGGTTGTTTGC CGGATCAAGAGCTACCAACTCTTTTCCGA 6306

QY 4871 aggtaaactggcttcagcagagcgagataccaaataactgtccttcttagtgagcagtagt 4930
|||||
Db 6305 AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTCTAGTGTAGCCGTAGT 6246

QY 4931 taggccaccacttcaagaactctgttagcaccgcctacatacctcgctcgtctaactcctgt 4990
|||||
Db 6245 TAGGCCACCACCTTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGCTAATCCTGT 6186

QY 4991 taccagtggtgctgctgccagtgggcgataaagtcgtgtcttaccggggttgacctcaagacgat 5050
|||||
Db 6185 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTGGACTCAAGACGAT 6126

QY 5051 agttaccgggataaaggcgagcggttcgggctgaacgggggttctgtgcacacagccagct 5110
|||||
Db 6125 AGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGTTCGTGCACACAGCCAGCT 6066

QY 5111 tggagcgaacgacctacacgaaactgagatacctacagctgagctatgagaaagcgcca 5170
|||||
Db 6065 TGGAGCGAACGACCTACACGAACTGAGATACCTACAGCTGAGCTATGAGAAAGCGCCA 6006

QY 5171 cgcttcccgaaggagaaaggcgagcagggtatccggttaagcggcagggctcggaacaggag 5230
|||||
Db 6005 CGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGGGTCCGAACAGGAG 5946

QY 5231 agcgcacgagggagcttccagggggaaacgcgcctgtatctttatagtcctgtcgggtttc 5290
|||||
Db 5945 ACGGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCTGTGCGGGTTC 5886

QY 5291 gccacctctgacttgagctcgatcttttacggttcctggccttttgcgcttttgcacaca 5410
|||||
Db 5885 GCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTACGGGTTCCTGGCCTTTTGTCTACA 5766

QY 5411 tgttcttccctgcgttatccctgattctgtggataaacggtattacccgctttgagtgag 5470
|||||
Db 5765 TGTCTTTCTCGCTTATCCCTTGATTTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG 5706

QY 5471 ctgataccgctcgcgcgagcgaacgacgagcgcgagcagtgagtgagtgaggaagcgg 5530
|||||
Db 5705 CTGATACCGCTCGCCGCAGCGAACGACCGAGCGCAGCGAGTCAGTGACGAGGAGCGG 5646

QY 5531 aagagcgcccaataacgcaaacccgcctctccccgcgcttgccggtattcattaatgcagct 5590
|||||
Db 5645 AAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGGCGGTGGCCGATTCTAATGACGCT 5586

QY 5591 ggcacgacaggtttccccgactggaaagcgggagtgagcgcaacgcaattaatgtgagtt 5650
|||||
Db 5585 GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 5526

QY 5651 agctcactcattaggccacccagggctttacactttatgcttccggctcgtatgtgtgtg 5710
|||||
Db 5525 AGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 5466

QY 5711 gaattgtgagcggataaactttcacacaggaacagctatgacctattacgcca 5767
|||||
Db 5465 GAATTGTGACGGGATAACAAATTTACACACAGGAAACAGCTATGACCATGATTACGCCA 5409

RESULT 8
AX114871/c
LOCUS AX114871 8153 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 31 from Patent WO0129208.
ACCESSION AX114871
VERSION AX114871.1 GI:14031813
KEYWORDS
SOURCE synthetic construct.

ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 8153)
AUTHORS Kuehn,R., von Melchener,H. and Altschmied,J.
TITLE Conditional gene trapping construct for the disruption of genes
JOURNAL Patent: WO 0129208-A 31 26-APR-2001;
ARTEMIS Pharmaceuticals GmbH (DE) ; Frankgen Biotechnologie AG (DE)
FEATURES
Location/Qualifiers
source 1..8153
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="vector PRK74"
BASE COUNT 1976 a 2133 c 2067 g 1977 t
ORIGIN

Query Match 45.7%; Score 2637; DB 6; Length 8153;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3131 cgaattcactggccgctgcttttacaaacgctgctgactgggaaaaaccctggcgttaccccaac 3190
|||||
Db 8136 CGAATTCACGTGGCCGTCGTTTACAAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCCAAC 8077

QY 3191 ttaatcgcccttgacgacacatccccctttcgccagctggcgtaaatagcgaagagggccgca 3250
|||||
Db 8076 TTAATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGGCCCCGCA 8017

QY 3251 cagatcgcccttcccaacagttgcgagcctgaatggcgaaatggcgctgatcggttatt 3310
|||||
Db 8016 CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCAATGGCGCTGATGCGGTATT 7957

QY 3311 ttctccttacgcatctgtgcggtatttcacacgcgcataatggtgcactctcagtacaatct 3370
|||||
Db 7956 TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGTCACACTCTCAGTACAATCT 7897

QY 3371 gctctgatccgcgcatagttaagccagccccgcacacccgccaaacacccgctgacgcgcct 3430
|||||
Db 7896 GCTCTGATGCCGCATAGTTAAAGCCAGCCCCGACACCCCGCAACACCCGCTGACCGGCCCT 7837

QY 3431 gacgggcttgcgtctccccgcgcacccgcatccgcttacagacaagctgtgacccgtctccgggagct 3490
|||||
Db 7836 GACGGGCTTGTCTGCTCCCCGCATCCGCTTACAGACAAGCTGTGACCGCTCTCCGGGAGCT 7777

QY 3491 gcatgtgtcagaggttttccacgctcatcacccgaaacgcgcgagacgaaaggccctcgtga 3550
|||||
Db 7776 GCATGTGTACAGAGGTTTTCACCGTCTATCACCGAAACCGCGCAGACGAAAGGCCCTCGTGA 7717

QY 3551 taagcctatttttataggttaagtgtcatgataataatggtttctttagacgtcaggtggca 3610
|||||
Db 7716 TAGCCCTATTTTATAGGTTAATGTCTATGATAATAATGGTTCTTAGACGTCAGGTGGCA 7657

QY 3611 ctttccggggaaatgtgcggaacccctatttgtttatttttctaaatacatattcaata 3670
|||||
Db 7656 CTTTTCGGGGAATGTGCGCGGAACCCCTATTGTGTTATTTTTCTAAATACATTCAATA 7597

QY 3671 tgtatccgctcatgagacaataaaccttgataaaatgcttcaataattgaaaaagggaaga 3730
|||||
Db 7596 TGATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA 7537

QY 3731 gtagagtattcaacatttccggtgcgccccttattcccccttttttgcggcattttgccttc 3790
|||||
Db 7536 GTATGAGTATTCAACATTTCCGTGTGCCCCCTTATTCCTCTTTTTCGGGCATTTTGGCCTTC 7477

QY 3791 ctgtttttgctcaccacgaaacgctggtgaaagtataaagatgctgaagatcagttgggtg 3850
|||||
Db 7476 CTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG 7417

QY 3851 cagagtggtgttacatcgaactggatctcaacacggtgaagatcccttgagagttttccgc 3910
|||||
Db 7416 CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTCGCC 7357

QY 3911 ccgaagaacgcttttcccaatgatgagcactttttaaagttctgctatgtggcgctattat 3970

|||||
Db 7356 CCGAAGAACGGTTTCCAAATGATGAGCACATTTAAAGTTCTGCTATGTGGCGGGTATTAT 7297
QY 3971 cccgtattgacgcgggcaagagcaactcggtcgccgcatacacactattctcagaatgact 4030
Db 7296 CCCGTATTGACGGCGGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAATGACT 7237
QY 4031 tgggtgagtactaccaggctcacagaaaaagcatcttacggatggcatgacagtaagagaat 4090
Db 7236 TGGTTGAGTACTCACCAAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA 7177
QY 4091 tatgcagtgctgccataaaccatgagtgataaacactcgggccaacttacttctgacaaacga 4150
Db 7176 TATGCAGTCTGCCATAACCATGAGTGATGAACACTGCGGCCAACTTACTTCTGACAAACGA 7117
QY 4151 tcggaggacccaaggagactaaaccgctttttgacaaacatggggagatcatgtaactcgcc 4210
Db 7116 TCGGAGGACCGAAGGAGCTAACCGCTTTTGTGCACAACATGGGGGATCATGTAACTCGCC 7057
QY 4211 ttgatcgttgggaaccggagctgaatgaagccatacacaacgacgagcgtgacaccacga 4270
Db 7056 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCGTGACACCACGA 6997
QY 4271 tgcctgtaggaatggcaacaactgttgcgaaaactattaaactggcgaaactacttactctag 4330
Db 6996 TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG 6937
QY 4331 ctccccggcaacaattaatagactggatggagcgcgataaaagtgcaggaccacttctgc 4390
Db 6936 CTTCCCGGCAACAATTAATAGACTGGATGGAGCGGATAAAGTTGAGGACCACTTCTGC 6877
QY 4391 gctcggcccttcgggctgctggtttattgctgataaaactctggaacgagcggtgagcgtgggt 4450
Db 6876 GCTCGGCCCTTCGGGCTGGCTGTTTATGTGTGATAAATCTGGAGCGGTGAGCGTGGGT 6817
QY 4451 ctcgcggtatacatgacgacactgggcccagatggtaagccctcccgatcgtagttatct 4510
Db 6816 CTCGGGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 6757
QY 4511 acacgacggggagtcaggcaactatcgatgaacgaaatagacagatcgctgagataggtg 4570
Db 6756 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 6697
QY 4571 cctcactgattaagcattggttaactgtcagaccaagtcttactcataatactttagattg 4630
Db 6696 CCTCACTGATTAAGCATTGGTAACGTGCAGACCAAGTTTACTCATATATACTTTAGATTG 6637
QY 4631 atttaaaacttcatttttaatttaaaggatctaggtgaagatccttttttgataaactca 4690
Db 6636 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 6577
QY 4691 tgaccaaatacccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaaaga 4750
Db 6576 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCCAGACCCCGTAGAAAAGA 6517
QY 4751 tcaaaggatcttcttgagatcctttttctgcgcgtaaatctgctgcttgcaaaacaaaaa 4810
Db 6516 TCAAAGGATCTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGTGCAAAACAAAA 6457
QY 4811 aaccacgctaccagcggtggtttgtttgccggatcaagagctaccacactcttttccga 4870
Db 6456 AACCACCGTACCAGCGGTGGTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGA 6397
QY 4871 aggtaactggcttcagcagagcgagatacacaataactgtccctcttagtgtagccgtagt 4930
Db 6396 AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTCTTAGTGTAGCCGTAGT 6337
QY 4931 taggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctaactcgt 4990
Db 6336 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTTCGCTCTGCTAATCCTGT 6277
QY 4991 taccagtggctgctgccagtgggcgataaagtcgtgtcttaccggggttgactcaagacgat 5050
|||||

Db 6276 TACCAGTGGCTGCTGCCAGTGGCGATAAAGTCGTGTCTTACGGGGTTGGACTCAAGACGAT 6217
QY 5051 agttaccgggataaggcgcagcgtcgggctgaacgggggttcgtgcacacagcccagct 5110
Db 6216 AGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT 6157
QY 5111 tggagcgaacgacctacacccgaactgagataacctacagctgagctatgagaaagcgcca 5170
Db 6156 TGGAGCGAAGACGACTACACCGAAGTACACCTGAGATACCTACAGCGTGAGTATGAGAAAGCGCCA 6097
QY 5171 cgcttcccgaaggagagaaaaggcggacaggtatccggtaagcggcgagggtcggaacaggag 5230
Db 6096 CGCTTCCGGAAGGAGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGTTCGGAACAGGAG 6037
QY 5231 agcgcacgagggagcttccagggggaaaacgcgcctggtatctttatagtcctgctcggttcc 5290
Db 6036 AGCGCACGAGGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCTCTGTCGGGTTTC 5977
QY 5291 gccacctgtgacttgagcgtcgatattttgtgatgctcgtcagggggcgagacctatgga 5350
Db 5976 GCCACCTCTGACTTGAGCGTCGATTTTGTGTGATGCTCGTCAGGGGGGGGAGCCTATGGA 5917
QY 5351 aaaaagcagcaacgcggcctttttacgggttccctggccttttgcggccttttgcacaca 5410
Db 5916 AAAACGCCAACAACGCGCCTTTTACGGTCTCTGGCCTTTTGTGCTGCACA 5857
QY 5411 tgtcttctcgtgcttatccccctgattctgtggataaacctgattaccgcctttgagtgag 5470
Db 5856 TGTTCTTCTGCTGCTTATCCCTGATTCTGTGGATAACCGTATTAACGCCTTTGAGTGAG 5797
QY 5471 ctgataccgctcgccgcagcgcgaacgcagcagcagcagcagcagcagcagcagcagcagc 5530
Db 5796 CTGATACCGCTCGCCGACGCGAACGACCGAGCGCAGCGAGTCACTGAGCGAGGAAGCGG 5737
QY 5531 aagagcgcccaatacgcgaacccgctctctcccgcgcggttgccgattcattaatgcagct 5590
Db 5736 AAGAGCGCCCAATACGCAAAACCCCTCTCCCCGCGCTTGGCCGATTCATTAATGCAGCT 5677
QY 5591 ggcacgacaggtttcccgactggaagcgggagcagtgagcgaacgcaataatgtgagtt 5650
Db 5676 GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTT 5617
QY 5651 agctcactcattaggcacccccaggctttacacacttttatgcttccggctcgtatgtgtg 5710
Db 5616 AGCTCACTCATTAGCACCCCGAGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGTG 5557
QY 5711 gaattgtgagcgggataaacaatttcacacaggaacacagctatgacctgattacgcca 5767
Db 5556 GAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCA 5500

RESULT 9

CVPGEN4Z/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

CVPGEN4Z 2746 bp DNA circular SYN 26-JAN-2000
Cloning vector pGEM-4Z.

X65305
X65305.2 GI:6782314
beta-lactamase; bla gene; cloning vector; lacZ gene; multiple

cloning site; promoter.
Cloning vector pGEM-4Z.
Cloning vector pGEM-4Z

artificial sequence; vectors.
1 (bases 1 to 2746)
Technical Services.

Direct Submission
Submitted (23-MAR-1992) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA

revised by [2]
2 (bases 1 to 2746)
Technical Services.

Direct Submission
Submitted (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA

QY 4324 actctagcttccggcgaacaattaatagactggatggagcggataaaagttgcaggacca 4383
Db 1536 ACTCTAGCTTCCCGCAACAATAATAGACTGGATGGAGCGGATAAAAGTTCAGGACCA 1477
QY 4384 cttctgcgctcgcccttccggctggctggtttattgctgataaaatctggagccggtgag 4443
Db 1476 CTTCTGGCTCGGCCCTTCCGGCTGGCTGATTGCTGATAAATCTGGAGCCGGTGAG 1417
QY 4444 cgtgggtctgcgggtatcatctgcagcactcggggccagatggttaagccctcccgatcgtat 4503
Db 1416 CGTGGTCTCGCGGTATCATTTGCAGCACTGGGCCCAGATGGTAAGCCCTCCCGTATCGTA 1357
QY 4504 gttatctacacgaggggagtcaggcaactatggtgaacgaataagacagatcgctgag 4563
Db 1356 GTTATCTACAGCGGGGAGTCAGGCACTATGGATGAACGAATAGACAGATCGCTGAG 1297
QY 4564 ataggtgcctcactgattaaagcattggttaactgtcagaccaagtttactcatatatactt 4623
Db 1296 ATAGGTGCCTCACTGATTAAAGCATTTGGTAAGTGTTCAGACCAAGTTTACTCATATATACTT 1237
QY 4624 tagattgatttaaaacttcattttttaatttaaaaggatctaggtgaagatcctttttgat 4683
Db 1236 TAGATTGATTAAAAACTTCATTTTAAATTAANAAGGATCTAGGTGAAGATCCTTTTGTAT 1177
QY 4684 aatctcatgaccaaatacccttaacgtgagtttctgttccactgagcgtcagacccccgta 4743
Db 1176 AATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTA 1117
QY 4744 gaaaagatcaaaaggatcttcttgagatcctttttctgcgcgttaactgcgtgcttgcaa 4803
Db 1116 GAAAAGATCAAAAGGATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGTGCAA 1057
QY 4804 acaaaaaaacccgcgtaccagcgggtgttggttggtccggatcaagagctaccaactctt 4863
Db 1056 ACAAAAAAACCCGCTACCAGCGGTGTTGTTGCCGGATCAAGAGTACCAACTCTT 997
QY 4864 tttccgaaggtaactggcttcagcagagcgagataccaaaatactgtccttcttagttag 4923
Db 996 TTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAG 937
QY 4924 ccgtagttagccaccacttcaagaactctgtagcaccgcctacatacctcgtctgcta 4983
Db 936 CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTACATACTCGCTCGCTGCTA 877
QY 4984 atcctgttacagtggtgctgctgcagtgagcagataagtcgtgtcttaccgggttgactca 5043
Db 876 ATCCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTGGACTCA 817
QY 5044 agacgatagtaccggtataagcgcagcggctcgggtgaacgggggttctgtagcacacag 5103
Db 816 AGACGATAGTTACCGGATAAAGCGCAGCGGTGGGCTGAACGGGGGTTCGTGCACACAG 757
QY 5104 ccagcttgagcgaacacgacctacacgaactgagatacctacagcgtgagctatgagaa 5163
Db 756 CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA 697
QY 5164 agcgccacgcttcccgaaggagaaagcggagacaggtatccggttaagcggcagggtcgga 5223
Db 696 AGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGA 637
QY 5224 acaggagagcgacagggagcttccagggggaaacgcctggtatctttatagtcctgtc 5283
Db 636 ACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCCTGTC 577
QY 5284 ggggttcgcacactctgacttgagcgtcgatttttgtgatgctgcgcagggggcggagc 5343
Db 576 GGGTTTCGCCACCTCTGACTTGAGCGTCCGATTTTGTGATGCTCGTCAGGGGGCGGAGC 517
QY 5344 ctatggaaaaacgcagcaacgcggcctttttacggttcctggccttttgcggcctttt 5403
Db 516 CTATGGAAAAACGCCAGCAACCGCGCCTTTTACGGTTCTCGGCCTTTTGTGCTGCCCTTTT 457
QY 5404 gctcacatgtttcttctgcgttatccctcgtattctgtggataaacgctattaccgccttt 5463

Db 456 GCTCACATGTTCTTCTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACGCCTTT 397
QY 5464 gagtgcgtgataccgctcgcgcagcgcgaacgagcgcagcagcagtcagtcagtcagcag 5523
Db 396 GAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCTCAGTGAGCGAG 337
QY 5524 gaagcgggaagcgcgccaataacgcaaacccctctccccgcgcgttgccgattcattaa 5583
Db 336 GAAGCGGAAGAGCGCCCAATACGCAACCCGCTCTCCCCGCGGTGGCCGATTCAATAA 277
QY 5584 tgcagctggcagcagaggtttccccgactggaaagcgggagtcagtcagtcagtcagtcag 5643
Db 276 TGCAGCTGGCACACACAGGTTTCCCAGACTGGAAGCGGCGCAGTGAGCGCAACGCAATAAT 217
QY 5644 gtgagttagctcaactcattagcaccacccaggttttacactttatgcttcgggctcgtatg 5703
Db 216 GTGAGTTAGCTCACTCATTTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATG 157
QY 5704 ttgtgtggaattgtgagcgggataaacaatttcacacaggaacagctatgaccatgattac 5763
Db 156 TTGTGTGGAATTGTGAGCGGATAACAATTTTACACAGGAACAGCTATGACCATGATTAC 97
QY 5764 gccca 5767
Db 96 GCCA 93
CVU47294 4392 bp DNA linear SYN 02-FEB-2000
Cloning vector pAdvantage<TM>, complete sequence.
U47294
U47294.2 GI:6854618
Cloning vector pAdvantage<TM>.
Cloning vector pAdvantage<TM>
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4392)
AUTHORS Groskreutz,D.J. and Schenborn,E.T.
TITLE Increased Gene Expression in Mammalian cell lines-using pAdvantage(TM) DNA as a Co-transfectant
JOURNAL Promega Notes 48, 8-12 (1994)
REFERENCE 2 (bases 1 to 4392)
AUTHORS Groskreutz,D.J. and Schenborn,E.T.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1996) D.J. Groskreutz, R&D, Promega Corporation, 5445 East Cheryl Parkway, Madison, WI 53711, USA
REFERENCE 3 (bases 1 to 4392)
AUTHORS Grooms,K.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) Technical Writing, Promega Corporation, 2800 Woodshollow Road, Madison, WI 53711-5399, USA
REMARK Sequence update by submitter
COMMENT On Feb 2, 2000 this sequence version replaced gi:1200457.
FEATURES
Location/Qualifiers
source 1..4392
/organism="Cloning vector pAdvantage<TM>"
/db_xref="taxon:45856"
misc_feature 1..1725
/note="adenovirus DNA insert"
misc_RNA 780..939
/product="adenovirus VAI RNA"
misc_RNA 1036..1198
/product="adenovirus VAI RNA"
gene complement(2904..3764)
/gene="Ampr"
CDS complement(2904..3764)
/gene="Ampr"
/note="ampicillin resistance"
/codon_start=1
/product="beta-lactamase"
/protein_id="AAA89081.1"

/db_xref="GI:1200458"
/translation="MSIQHFRVALIPFFAAFLPVFAPHETLVKVKDAEDQLGARVGY
IELDLNSGKILESFRPEERFPMSTFKVLICGAVLSRIDAGQEQLGRRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGGPKELTAFIHNMGDHSVTRL
DRWPELNEAIPNDERDTHMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
LRSALPAGWFIAIDKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
EIGASLIKHW"
BASE COUNT 973 a 1177 c 1265 g 977 t
ORIGIN

Query Match 45.7%; Score 2635.4; DB 12; Length 4392;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2636; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	3131	cgaattcactggccgctcggttttacaaagtcgtgactgggaaaaaccctggcgttaccacaac	3190
Db	4366	CGAATTCACTGGCCGTCGTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAAC	4307
QY	3191	ttaatgcgcttgacagacacatccccctttcgccagctggcgtaataagcggaagagggcccgca	3250
Db	4306	TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGCCCGCA	4247
QY	3251	ccgatacgcccttcccaacagttgcgcagcctgaatggcgaaatggcgccctgatcgcggtatt	3310
Db	4246	CCGATCGCCCTTCCCAACAGTTGGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATT	4187
QY	3311	ttctccttacgcatctgtgcggtatttcacaccgcataatggtgcactctcagtacaatct	3370
Db	4186	TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGTTGCACCTCTCAGTACAATCT	4127
QY	3371	gctctgatgccgcatagttaagccagccccgcacacccgcgaacaccccgctgacgcgcct	3430
Db	4126	GCTCTGATGCCGATAGTTAAGCCAGCCCCGACACCCCGCAACACCCGCTGACGCGCCCT	4067
QY	3431	gacgggcttgctgctgcctcccgcatccgcttacagacaagctgtgacccgtctccgggagct	3490
Db	4066	GACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT	4007
QY	3491	gcattgctcagagggttttcaccgctcatcaccgaaaacgcggagacgaaaaggccctcgtga	3550
Db	4006	GCATGTGTCAGAGGTTTTCACCGTTCATCACCGAAACCGCGGAGACGAAAGGGCCTCGTGA	3947
QY	3551	taagcctatttttataggttaatgtcatgataataatggtttctttagacgtcaggtggca	3610
Db	3946	TACGCCTATTTTATAGGTTAATGTTCATGATAATAATGGTTTCTTAGACGTGAGGTGGCA	3887
QY	3611	cttttcggggaaatgtgcgcggaacccctatttggtttatttttctaaatacattcaaaata	3670
Db	3886	C'TTTTCGGGGAAATGTGCGCGGAACCCCTATTGTGTTATT'TTCTAAATACATTCAAATA	3827
QY	3671	tgtatccgctcatgagacaaataaaccttgataaaatgcttcaataatattgaaaaagggaaga	3730
Db	3826	TGTATCCGCTCATGAGACAAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA	3767
QY	3731	gtatgagtattcaacatttcggtgcgccccttattcccttttttgcggcattttgccttc	3790
Db	3766	GTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTT'TTTCGGGCATTTTGCCTTC	3707
QY	3791	ctggttttgcctcaccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg	3850
Db	3706	CTGTTT'TTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTG	3647
QY	3851	cacagtggtgtacatcgaaactggatctcaacagcggtaagatccttgagagttttcgcc	3910
Db	3646	CACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC	3587
QY	3911	ccgaagaacggttttcccaatgatgagcaccttttaaagttcgtctatgtggcgcggtattat	3970
Db	3586	CCGAAGAACGTTT'TCCAATGATGAGCAC'TTTTAAAGTTCGTATGTGGCGCGGTATTAT	3527
QY	3971	cccgatttgacgcgggcaagaggaactcggtcgcgcgcaatacactattctcagaatgact	4030

Db	3526	CCCGTATTGACGCCGGGCAAGAGCAACTCGGTCCGCCCATACACTATTCTCAGAAATGACT	3467
QY	4031	tggttgagtactcaccagtcacagaaaaagcatctttacggatggcatgacagtaagagaaat	4090
Db	3466	TGGTTCAGTACTCACCAGTCAACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAAT	3407
QY	4091	tatgcagtgcgccataaaccatgagtgataaacactgcggccaaacttacttctgacaaacga	4150
Db	3406	TATGCAGTGTGCCATAACCATGAGTGATAAACACTCGCGGCAACTTACTTCTGACAAACGA	3347
QY	4151	tcggaggaccggaaggagctaaaccgttttttgcacaacatgggggatcatgtaactcgcc	4210
Db	3346	TCGGAGGACCGGAAGGAGCTAACCGCTTTT'TTGCACAACATGGGGGATCATGTAACTCGCC	3287
QY	4211	ttgatcgcttgggaaccggagctgaatgaagccataaccaaaacgacgagcgtgacacccacga	4270
Db	3286	TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCCACGA	3227
QY	4271	tgcctgtagcaatggcacaacacgttgcgcaaaactattaaactggcggaactacttactctag	4330
Db	3226	TGCCCTGTAGCAATGGCAACACGTTGGCAAACTATTAACTGGCGAACTACTTACTCTAG	3167
QY	4331	cttcgcgggcaacaattaatagactcggtgggttattgctgataaaatcggagccggtgagcgtgggt	4390
Db	3166	C'TTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACCTTCTGC	3107
QY	4391	gctcgcccttccggctgggtggttattgctgataaaatcggagccggtgagcgtgggt	4450
Db	3106	GCTCGGCCCTTCCGGCTGGCTGGT'TTATTGCTGATAAACTCTGAGCCCGGTGAGCGTGGGT	3047
QY	4451	ctcgcggtatcatctgcagcactggggccagatggtaagccctcccgatatcgtagttatct	4510
Db	3046	CTCGGGTATCATTTGCAGCACTGGGCCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT	2987
QY	4511	acagacggggagtcaggccaactatggatgaacgaaatagacagatcgctgagataggtg	4570
Db	2986	ACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG	2927
QY	4571	cctcactgattaaagcatttggttaactgtcagaccaagtttactatatactatttagattg	4630
Db	2926	CCTCACTGATTAAAGCATTTGGTAAC'TGTCAGACCAAGTTTACTCATATATACTTTAGATTG	2867
QY	4631	atttaaaacttcatttttttaatttaaaaggatctaggtgaagatcccttttggataatctca	4690
Db	2866	ATTTAAAACTTCATTTT'TAATTTAAAGGATCTAGGTGAAGATCCTTTT'TTGATAATCTCA	2807
QY	4691	tgacaaaaatcccttaacgtgagtttcttcactgagcagaccccgtagaaaaaga	4750
Db	2806	TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAGA	2747
QY	4751	tcaaaggatctcttgagatcccttttttctgcgcgtaatctgctgcttgcaaacaaaaa	4810
Db	2746	TCAAAGGATCTTCTTGAGATCC'TTTTCTGCGCGTAACTCTGCTGCTTGCAAAACAATAA	2687
QY	4811	aaccacgcctaccagcgggtgttggtttgcgggatcaagagctaccaactctttttccga	4870
Db	2686	AACCACCGCTACCGCGGTGGTTGT'TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGA	2627
QY	4871	aggtaaactggcttcagcagagcgcataccaaaactgtccttctagtgtagccgtagt	4930
Db	2626	AGGTAAC'TGGCTTCAGCAGAGCGCAGATACCAAA'TACTGT'TCTTAGTGTAGCCG'TAGT	2567
QY	4931	taggccaccacttcaagaactctgtagcaccgcctacatactcgtctctgtaatcctgt	4990
Db	2566	TAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGTAAATCCCTGT	2507
QY	4991	taccagtggtgctgctgccagtgccgataaagtcgtgtcttacggggttggactcaagacgat	5050
Db	2506	TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTTGGACTCAAGACGAT	2447
QY	5051	agttaccgggataaaggcgacggtcgggctgaacggggggttcgctgcacacagccccagct	5110
Db	2446	AGTTACCGGATTAAGCGCACCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCCAGCT	2387

Db 2182 TGACGGGCTTGCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGC 2241

QY 3490 tgcattgtgtcagaggtttttaccggtcatcacccgaaacgcgcgagacgaaaaagggcctcgtg 3549

Db 2242 TGCATGTGTcAGAGGTTTTACCGTcATCACCGAAACCGCGGAGACGAAAGGCCCTCGTG 2301

QY 3550 ataagcctatttttataggttaattatgtcatgataataatggtttcttagacgtcaggtggc 3609

Db 2302 ATAGCCCTATTTTTATAGGTTAATGTcATGATAATAATGGTTCTTAGACGTcAGGTGGC 2361

QY 3610 acttttcggggaattgtcgcgggaacccctatttgtttatttttctaaatacatcfaat 3669

Db 2362 ACTTTTCGGGGAATGTGCGGGAACCCCTATTtGTtATTtTCTAAATAcATTCAAAT 2421

QY 3670 atgtatccgctcatgagacaataacccctgataaaatgcttcaataataattgaaaaagggaag 3729

Db 2422 ATGTATCCGCTCATGAGACAATAAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAG 2481

QY 3730 agtatgagtattcaacatttcggtgtgcgcccatttcccttattcccttttttgcggcatittgcctt 3789

Db 2482 AGTATGAGTATTCACATTTCCGTGTGCCCCtTATtCCCCtTTtTTTGCGGCATtTTGCCtT 2541

QY 3790 cctgtttttgctcaccaccagaaacgctggtgaaagttaaagatgctgaagatcagttcgggt 3849

Db 2542 CCTGTtTTTGCTCACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGT 2601

QY 3850 gcacgagtgggtttacatcgaaactggatctcacaacgcggtaagatcccttgagagttttcgc 3909

Db 2602 GCACGAGTGGGTtACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTtTTCG 2661

QY 3910 cccgaagaacgttttcccaatgatgagcaccttttaaagttctgctatgtggcgcggtatta 3969

Db 2662 CCCGAAGAACGTtTTCCAATGATGAGCACtTTTAAAGTtTCTGCTATGTGGCGCGTATTA 2721

QY 3970 tcccgattgacgcgggcaagagcaactcggtcgcgcgcatacacactattctcagaatgac 4029

Db 2722 TCCCGTATTGACGCGGGCAAGAGCAACTCGGTGCGCGCATACACTATtTCTCAGAATGAC 2781

QY 4030 ttggttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaa 4089

Db 2782 TTGGTGTAGTACTCACcAGTcACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA 2841

QY 4090 ttatgcagtgctgccataaacatgagtataaacactgcggccaacttacttctgacaaag 4149

Db 2842 TTATGCAGTGCTGCCATAACCATGAGTGATAACACTCGCGGCAACTTACTtCTGACAAcG 2901

QY 4150 atcggaggaccgaaggagctaaccgcttttttgacaaacatgggggatcatgtaaactcgc 4209

Db 2902 ATCGGAGGACCGGAAGGAGCTAACCGCTTTtTTGCACAACATGGGGGATCATGTAACTCGC 2961

QY 4210 cttgatcgttgggaaccggagctgaatgaagccataaccaacgacgagcgtgacaccacg 4269

Db 2962 CTTGATCGTtGGGAACCGGAGCTGAATGAAGCCATACCAAAcGACGAGCGTGACACCCACG 3021

QY 4270 atgcctgtagcaatggcaacaacgcttgcgcaaaactattaactggcgaaactactactcta 4329

Db 3022 ATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATtAACTGGCGAACTACTTACTCTA 3081

QY 4330 gcttcccggaacaattaatagactggatggaggcgggataaaagtgcaggaccacttctg 4389

Db 3082 GCtTCCCGGCAACAATTAATAGACTGGATGGAGGGGGATAAAAGTTGCAGGACCACTtCTG 3141

QY 4390 cgctcggcccttccggctggctggtttattgctgataaaatctggagccggtgagcgtggg 4449

Db 3142 CGTcGGCCCCtTCCGGCTGGCTGGTtTATTGCTGATAAACTTGGAGCCCGGTGAGCGTGGG 3201

QY 4450 tctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatc 4509

Db 3202 TCTCGGGTATCATtGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTtATC 3261

QY 4510 tacagcagcgggagtcaggcaactatggatgaacgaaatagacagatcgcctgagataggt 4569

Db 3262 TACACGACGGGGAGTCAGGCCAACTATGGATGAACGGAATAGACAGATCGCTGAGATAGGT 3321

QY 4570 gcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagatt 4629

Db 3322 GCCtCACTGATTAAGCAtTGGTAActGTcAGACCAAGtTtTACTATATATAcTtTtTAGATt 3381

QY 4630 gatttaaaaacttcattttttaatttaaaaggatctaggtgaagatccctttttgataatctc 4689

Db 3382 GATTtAAAAACTTCATtTTTAATtTAAAAAGGATCTAGGTGAAGATCCTtTTTtTGATAATCTC 3441

QY 4690 atgaccaaaaatcccttaacgtgagtttctgctccactgagcgcagacccccgtagaaaaag 4749

Db 3442 ATGACCAAAATCCCTTAACGTGAGTtTTCGTTCCACTGAGCGGTcAGACCCCGTAGAAAAG 3501

QY 4750 atcaaaaggatcttcttgagatcccttttctgcgcgtaatactgcgtctgcgtaacaaaa 4809

Db 3502 ATCAAAAGGATCTTCTTGAGATCCCTTTtTtTCTGCGGTAAATcGTGCTTGCAAAACAAA 3561

QY 4810 aaaccaccgctaccagcgggtgttgttggccggatcaagagctaccaactcttttccg 4869

Db 3562 AAACCACCGCTACCAGCGGTGGTtTGTtTGCCGGATCAAGAGCTACCAACTCTtTTtTCCG 3621

QY 4870 aaggtaaactggcttcagcagagcgagataccaaatactgtccttctagtgtagccgtag 4929

Db 3622 AAGGTAACTGGCTTCAGCAGAGCGAGATACCAAACTAcTGTCTCTTAGTGTAGCCGTAG 3681

QY 4930 ttaggccaccacttcaagaactctgtagcaccgcctacatacacctcgcctcgtctaatacctg 4989

Db 3682 TTAGSCCACCACtTCAAGAACTCTGTAGCACCCGCTACATACCTCGCTCTGTAAATCCTG 3741

QY 4990 ttaccagtggtcgtgcagtgccagtaagtcgctgtcttaccggggttgactcaagacga 5049

Db 3742 TTACCAGTGGCTGTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTGGACTCAAGACGA 3801

QY 5050 tagttaccggataaaggcgcagcgtcggtcggtgaacgggggttcgtgcacacagcccagc 5109

Db 3802 TAGTTACCGGATAAAGCGCGACGCTCGGGCTGAACGGGGGTTCGTGCACACAGCCCAGC 3861

QY 5110 ttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgaaaagcgc 5169

Db 3862 TTGGAGCGAAACGACCTACACCCGAActGAGATACCTACAGCGTGAGCTATGAGAAAGCGCC 3921

QY 5170 acgcttcccgaaggagaaaaggcgcagaggtatccggtaagcgcagggctcggaacagga 5229

Db 3922 ACGCTTCCCGAAGGGAGAAAGCGCGACAGGTATCCGGTAAGCGCAGGGTcGGAACAGGA 3981

QY 5230 gagcgcagcagggagcgttcccaggggaaaaacgcctgggtatctttatagtcctgctgggttt 5289

Db 3982 GAGCGCAGGAGGAGCTTCCAGGGGGAAACGCCCTGGTATCTtTATAGTCTCTGTcGGGTTT 4041

QY 5290 cgccaactctgacttgagcgtcgatttttgatgctcgtcaggggggcgagcctatgg 5349

Db 4042 CGCCACCTCTGACTTGAGCGTCGATtTTTGATGCTCGTCAGGGGGCGGAGCCTATGG 4101

QY 5350 aaaaacgcagcaaacgcgcctttttacgggttcctggccttttgctggccttttgctcac 5409

Db 4102 AAAAACGCCAGCAACGCGGCTTTTtACGGTTCCTGGCCTTTtTGTGGCCTTTtTGTCTCAC 4161

QY 5410 atgttcttccctgcgttatccctgattctgtggataacgctattaccgcctttgagtgga 5469

Db 4162 ATGTCTCTTCTCGGTTATCCCCCTGATtCTGTGGATAACCGTATtACCGCCTTTtTGAGTGA 4221

QY 5470 gctgataccgctcgcgcgcagccgaacgcagcgcagcgcagtcagtcagtcaggaagcg 5529

Db 4222 GCTGATACCGCTCGCCCGCAGCCGAACGACCGAGCGCAGAGTcAGTGAGCGGAGGAAGCG 4281

QY 5530 gaagcgcgccccaaatacgaaaaacgcctctccccgcgcgttgccgattcattaatgcagc 5589

Db 4282 GAAGAGCGCCCCAAATACGCAAAACCGCTCTCCCGCGGTtGGCGGATTCATTAATGCAGC 4341

QY 5590 tggcagcacaggtttccccgactgaaaagcgggcagtcgagcgaacgcgaattaatgtgagt 5649

Db 4342 TGGCAGCACAGGTTTCCCGACTGGAAGAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGT 4401

```
Qy 5650 tagctcaactaattagggcaccacccaggctttacacactttatgcttccggctcgatgtgtgt 5709
|||||
Db 4402 TAGCTCACTCATAGGCACCCAGGCTTACACTTTATGCTTCCGGCTCGTATGTTGTGT 4461
|||||
Qy 5710 ggaattgtgagggataacaatttcacacaggaaacagctatgaccatgattacgcc 5766
|||||
Db 4462 GGAATTGTGAGCGGATAACAATTTCACACAGGAACAGCTATGACCATGATTACGCC 4518
|||||

RESULT 12
U02456
LOCUS U02456 4519 bp DNA circular SYN 29-MAR-1996
DEFINITION Cloning vector pRAJ275, complete sequence.
ACCESSION U02456
VERSION U02456.1 GI:413822
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Kitts,P.A.
TITLE CLONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4519)
AUTHORS Jefferson,R.A.
TITLE Plant reporter genes: the GUS gene fusion system
JOURNAL (in) Setlow,J.K. (Ed.);
GENETIC ENGINEERING: PRINCIPLES AND METHODS;
10: 247-263 (1988)
REFERENCE 3 (bases 1 to 4519)
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH. This vector is no longer
available from CLONTECH and CLONTECH will not update or revise this
sequence.
FEATURES
source Location/Qualifiers
1. .4519
/organism="Cloning vector pRAJ275"
/db_xref="taxon:31825"
BASE COUNT 1130 a 1131 c 1190 g 1068 t
ORIGIN

Query Match 45.7%; Score 2635.4; DB 12; Length 4519;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2636; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3130 gcgaattcactggccgctcgctttttacaacgctgtagctgggaaaccctggcgttacccaa 3189
|||||
Db 1883 GGGAAATTCAGTGGCCGCTCGTTTTACAACGCTGCTGACTGGGAAACCCCTGGCGTTACCCAA 1942
|||||
Qy 3190 cttaatcgccctgcagcacatcccccttcgccagctggcgtaataagcgaagagggccgc 3249
|||||
Db 1943 CTTAATCGCCTTGCAGCACATFCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGC 2002
|||||
Qy 3250 accgatcgcccttcccaacagtgtgcgagcctgaatggcgaaatggcgccctgatgcggtat 3309
|||||
Db 2003 ACCGATCGCCCTTCCCAACAGTTGCGAGCCTGAATGGCGAATGGCGCCTGATGCGGTAT 2062
|||||
Qy 3310 ttctctctacgcatctgtgcggtatttcacaccgcataatggtcacactctcagtacaatc 3369
|||||
Db 2063 TTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACTCTCAGTACAATC 2122
|||||
Qy 3370 tgctctgatccgcatagttaagccagccccgcacacccgccaaaccccgctgacgcgcc 3429
|||||
Db 2123 TGCTCTGATGCGGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCC 2182
|||||
Qy 3430 tgacgggcttgctgctcccgccatcccgcttacagacaagctgtgaccgtctccgggagc 3489
|||||
```

```
Db 2183 TGACGGGCTTGCTGTCTCCCGGCATCCGCTTACAGACAAGCTGTGACCCTCTCCGGGAGC 2242
|||||
Qy 3490 tgcattgtcagaggtttttcacccgtcatcaccgaaacgcgcgagacgaaaggcctcgtg 3549
|||||
Db 2243 TGCATGTGTCAAGAGGTTTTCACCGCTCATCCGAAACGCGCGAGACGAAAGGCCCTCGTG 2302
|||||
Qy 3550 atacgcctatttttataggttaattatgcatgataataatggtttcttagacgctcagggtggc 3609
|||||
Db 2303 ATACGCCCTATTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTAGACGTCAGGTGGC 2362
|||||
Qy 3610 acttttcgggggaaatgtgcgcggaacccctattttgtttattttttaaataacattcaaat 3669
|||||
Db 2363 ACTTTTCGGGGAATGTGCGGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAAT 2422
|||||
Qy 3670 atgtatccgctcatgagacaataaacccctgataaaatgcttcaataatattgaaaaaggaag 3729
|||||
Db 2423 ATGTATCCGCTCATGAGACAATAACCCCTGATAAAATGCTTCAATAATATTGAAAAAGGAAG 2482
|||||
Qy 3730 agtatgagtattcaacatttccgctgctgcgcccttattcccttttttgcggccattttgcctt 3789
|||||
Db 2483 AGTATGAGTATTCAACATTTCCGCTGTCCGCTTATTCCCTTTTTCGGGCATTTTGCCCTT 2542
|||||
Qy 3790 cctgtttttgctcaccaccagaaacgcctggtgaaagtaaaagatgctgaagatcagtttggtg 3849
|||||
Db 2543 CCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGT 2602
|||||
Qy 3850 gcacgagtgggttacatcgaaactggatcgaacagcggtgaagatccttgagagtttctgc 3909
|||||
Db 2603. GCACGAGTGGGTTACATCGAACTGGATGCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGC 2662
|||||
Qy 3910 cccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcggtatta 3969
|||||
Db 2663 CCCGAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTA 2722
|||||
Qy 3970 tcccgtattgacgcgggcaagagcaactcggtcgccgcatacacactattctcagaatgac 4029
|||||
Db 2723 TCCCGTATTGACGCGCGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAAATGAC 2782
|||||
Qy 4030 ttggttgagtactcaccagtcacagaaaaagcatcttacggatggcatgacagtaagagaa 4089
|||||
Db 2783 TTGCTTGAGTACTCACCAGTCAACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAA 2842
|||||
Qy 4090 ttatgcagtgtgccataaacatgagtgataaacactcgggccaacttacttctgacaaacg 4149
|||||
Db 2843 TTATGCAGTGTGCCATAACCATGAGTGATAACTGCGGCCAACTTACTTCTGACAAACG 2902
|||||
Qy 4150 atcgggagaccgaaggagctaaaccgttttttgcaacaacatgggggatcatgtaactgc 4209
|||||
Db 2903 ATCGGAGACCGGAAGGAGCTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGC 2962
|||||
Qy 4210 cttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacg 4269
|||||
Db 2963 CTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACG 3022
|||||
Qy 4270 atgccttagcaatggcaacaacgttgcgcaaaactattaaactgggaactacttactcta 4329
|||||
Db 3023 ATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTA 3082
|||||
Qy 4330 gcttcccgcccttccggctggctggtttatttgctgataaaatcctggagccggtgagcgtggg 4389
|||||
Db 3083 GCTTCCCGGCAACAATTAATAGACTGGATGGAGGGGGATAAAGTTGCAGGACCACCTTCTG 3142
|||||
Qy 4390 cgctcgcccttccggctggctggtttatttgctgataaaatcctggagccggtgagcgtggg 4449
|||||
Db 3143 CGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAAATCTGAGCCCGGTGAGCGTGGG 3202
|||||
Qy 4450 tctcgcggtatcatgtgcagcactggggccagatggtaagccctcccgtagtagttatc 4509
|||||
Db 3203 TCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATC 3262
|||||
Qy 4510 tacacgacggggagtcaggcaactatgtagaacgaaatagacagatcgctgagataggt 4569
|||||
Db 3263 TACACGACGGGGAGTCAGGCCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGT 3322
|||||
```

QY 4570 gcctcaactgattaagcatgggtaaactgtcagaccaaagtttactcatatatactttagatt 4629
|||||
Db 3323 GCCTCACTGATTAAAGCATTGGTAACGTGTAGACCAAGTTTACTCATATATACTTTAGATT 3382

QY 4630 gatttaaaacttcatttttaatttaaaggatctaggtgaagatccttttttgataatctc 4689
|||||
Db 3383 GATTTAANAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTC 3442

QY 4690 atgaccaaatacccttaacgtgagtttctgctccactgagcgtcagaccccgtagaaaaag 4749
|||||
Db 3443 ATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAAG 3502

QY 4750 atcaaaaggatcttcttgagatcccttttttctgcgcgtaaatctgctgcttgcaaacaaaa 4809
|||||
Db 3503 ATCAAAAGGATCTTCTTGAGATCCTTTTCTGCGCGTAATCTGCTGTGCAAAACAAAA 3562

QY 4810 aaaccaccgctaccagcggtyggtttgtttgcccggatcaagagctaccacactctttttccg 4869
|||||
Db 3563 AAACCACCGCTACCAGCGGTGGTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCGG 3622

QY 4870 aaggtaactggcttcagcagcgcagatataccaaaatactgctcccttctagtgtagccgtag 4929
|||||
Db 3623 AAGGTAACCTGGCTTCAGCAGACGCGCAGATACCAATACTGTCTCTTCTAGTGTAGCCGTAG 3682

QY 4930 ttaggccaccactcaagaactctgtagcacccgcctacataccctcgctctgctaactctg 4989
|||||
Db 3683 TTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGTCTAATCCTG 3742

QY 4990 ttaccagtggtgctgctccagtgccgataaagtctgtcttaccggggttgactcaagacga 5049
|||||
Db 3743 TTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCGTCTTACCGGGTTGGACTCAAGACGA 3802

QY 5050 tagttaccgggataaaggcgagcggctcgggctgaacggggggttcgtgcacacagcccaagc 5109
|||||
Db 3803 TAGTTACCGGATAAGGCGCAGCGGTTCGGCTGAACGGGGGGTTCGTGCACACAGCCCAAGC 3862

QY 5110 ttggagcgaaacgacctacacggaactgagataacctacagcgtgagctatgagaaaagcgcc 5169
|||||
Db 3863 TTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCC 3922

QY 5170 acgcttccccgaaggagaaagcggacaggtatccggttaagcggcagggtcggaaacagga 5229
|||||
Db 3923 ACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGA 3982

QY 5230 gagcgcacgaggagcttcagggggggaacgcctggtatctttatagtcctgctcggyttt 5289
|||||
Db 3983 GAGCGCACGAGGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGTCTCTGTCGGGTTT 4042

QY 5290 cgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggagcgcctatgg 5349
|||||
Db 4043 CGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCGAGGGGGCGGAGCCTATGG 4102

QY 5350 aaaaacgcccagcaacggccctttttacgggttcctggcccttttgctggccttttgctcac 5409
|||||
Db 4103 AAAAAGCCAGCAACGGCGCCTTTTACGGTTCCTGGCCCTTTTGCTGGCCTTTTGTCTCAC 4162

QY 5410 atgttcttctcctgcgttatccctgattctctgtggataaacggtattaccgccttttgagtga 5469
|||||
Db 4163 ATGTTCTTTCTTGCCTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTTGAGTGA 4222

QY 5470 gctgataccgctcgccgcagcgaacgacgcgagcgcagcagtcagtcagtcagtcagtcagtc 5529
|||||
Db 4223 GCTGATACCGCTCGCCCGCAGCGCAACGACCGGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG 4282

QY 5530 gaagagcgcccaatacgcgaacacgcgcctctccccgcgcgttggccgatttcattaatgcagc 5589
|||||
Db 4283 GAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCCGCGGTGGCCGATTTCATTAATGCAGC 4342

QY 5590 tggcagcacaggtttccccgactggaaaagcgggagtcgagcgcaacgcaataatgtgagt 5649
|||||
Db 4343 TGGCAGCACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGT 4402

QY 5650 tagtcactcattagggcaccacgggcttttacactttatgcttccggctcgtatgttgt 5709
|||||
Db 4403 TAGTCACTCATTAGGCACCCCGAGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGT 4462

QY 5710 ggaattgtgagcggataacaatttcacacaggaacagctatgaccatgattacgcc 5766
|||||
Db 4463 GGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCC 4519

RESULT 13
XXU02437
LOCUS
DEFINITION Cloning vector pT3/T7-LUC, 4674 bp DNA circular SYN 29-JAN-1997
ACCESSION U02437
VERSION U02437.1 GI:413803
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4674)
AUTHORS Kitts,P.A.
TITLE Cloning vector pT3/T7-LUC.
JOURNAL Cloning vector pT3/T7-LUC
REFERENCE 2 (bases 1 to 4674)
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This vector can be obtained from CLONTECH Laboratories, Inc., 1020
East Meadow Circle, Palo Alto, CA 94303, USA. To place an order
call (415) 424-8222 or (800) 662-2566, extension 1. International
customers, please contact your local distributor. For technical
information, call (415) 424- 8222 or (800) 662-2566, extension 3.
This sequence has been compiled from information in the sequence
databases, published literature and other sources, together with
partial sequences obtained by CLONTECH; this vector has not been
completely sequenced. If you suspect there is an error in this
sequence, please contact CLONTECH's Technical Service Department at
(415) 424-8222 or (800) 662-2566, extension 3 or E-mail
TECH@CLONTECH.COM.

FEATURES
source
1. .4674
/organism="Cloning vector pT3/T7-LUC"
/db_xref="taxon:31848"
BASE COUNT 1239 a 1092 c 1141 g 1202 t
ORIGIN

Query Match 45.7%; Score 2635; DB 12; Length 4674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3133 aattcaactggccgctgctttttacaagctgctgactgggaaaaacccctggcttaccacactt 3192
|||||
Db 2008 AATTCACTGGCCGTCGTTTACAACTGCTGACTGGGAAACCCCTGGCGTTACCCAACTT 2067

QY 3193 aatcgcccttgacgacacatcccccttttcgacagctggcgtaataagcgaagagcccgacc 3252
|||||
Db 2068 AATCGCCTTGCAGCACATCCCCCTTTTCGCCAGCTGGCGGTAAATAGCAAGAGCCCCGCACC 2127

QY 3253 gatcgcccttcccaaacagttgcgagcctgaaatggcgaaatggcgccctgatcggtattttt 3312
|||||
Db 2128 GATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTT 2187

QY 3313 ctccctacgcatctgtgcggtattttcacaccgcgatatggtgcactctcagtacaatctgc 3372
|||||
Db 2188 CTCCTTAGCGCATCTGTGCGGTATTTCACACCGCATATGTTGCTACTCTCAGTACAATCTGC 2247

QY 3373 tctgatcgccgcatagttaagccagccccgcacacccgcgaaccccgctgacgcgccctga 3432
|||||
Db 2248 TCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA 2307

QY 3433 cgggcttgtctgctccccggcatccgcttacagacaagctgtaccgctctcccggaagctgc 3492

Db 2308 C G G G C T T G T C T C C G G C A T C C G C T T A C A G A C A A G C T G T G A C C G T C T C C G G G A G C T G C 2367
Qy 3493 a t g t g t c a g a g g t t t c a c c g t c a t c a c g a a a c g c g a g a c a a g g c c t c g t g a t a 3552
Db 2368 A T G T G T C A G A G G T T T C A C C G T C A T C A C G A A C G C G G A G A C G A A A G G C C T C G T G A T A 2427
Qy 3553 c g c c t a t t t t a t a g g t t a a t g t c a t g a t a a a t a a t g g t t t c t t a g a c g t c a g g t g g c a c t 3612
Db 2428 C G C C T A T T T T A T A G G T T A A T G T C A T A T A A T A A T G G T T T C T T A G A C G T C A G G T G G C A C T 2487
Qy 3613 t t t c g g g g a a a t g t g c g c g g a a c c c t a t t t g t t t a t t t t t c t a a a t a c a t t c a a a t a t g 3672
Db 2488 T T T C G G G G A A A T G T G C G C G A A C C C C T A T T T G T T A T T T T T C T A A A T A C A T T C A A A T A T G 2547
Qy 3673 t a t c c g t c a t g a g a c a a t a a c c c t g a t a a a t g c t t c a a t a a t a t t g a a a a g g a a g a g t 3732
Db 2548 T A T C C G C T C A T G A G A C A A T A A C C C T G A T A A A T G C T T C A A T A A T A T T G A A A A G G A A G A G T 2607
Qy 3733 a t g a g t a t t c a a c a t t t c c g t g t c g c c t t a t t c c c t t t t t t g c g c a t t t t g c c t t c o t 3792
Db 2608 A T G A G T A T T C A A C A T T T C C G T G T C G C C C T A T T C C C T T T T T T G C G C A T T T G C C T T C C T 2667
Qy 3793 g t t t t t g t c a c c c a g a a a c g t g t g a a a g t a a a a g a t g c t g a a g a t c a g t t g g g t g c a 3852
Db 2668 G T T T T T G C T C A C C C A G A A C G C T G G T G A A A G T A A A A G A T G C T G A A G A T C A G T T G G G T G C A 2727
Qy 3853 c g a g t g g g t t a c a t c g a a c t g g a t c t c a a c a g c g g t a a g a t c c t t g a g a g t t t t c g c c c c 3912
Db 2728 C G A G T G G G T T A C A T C G A A C T G G A T C T C A A C A G C G G T A A G A T C C T T G A G A G T T T T C G C C C C 2787
Qy 3913 g a a g a a c g t t t t c c a a t g a t g a g c a c t t t t a a g t t c t g c a t g t g g c g c g g t a t t a t c 3972
Db 2788 G A A G A A C G T T T T C C A A T G A T G A G C A C T T T T A A G T T C T G C T A T G T G G C G G G T A T T A T C C 2847
Qy 3973 c g t a t t g a c g c c g g g c a a g a g a c a a c t c g g t c g c g c a t a c a c t a t t c t c a g a a t g a c t t g 4032
Db 2848 C G T A T T G A C G C C G G G C A A G A G A C A C T C G G T C G C C G C A T A C A C T A T T C T C A G A A T G A C T T G 2907
Qy 4033 g t t g a t a c t c a c c a g t c a c a g a a a a g c a t c t t a c g g a t g g c a a c t a c t t c t g a c a a c g a t c 4092
Db 2908 G T T G A G T A C T C A C C A G T C A C A G A A A A G C A T C T T A C G G A T G G C A T G A C A G A T A A 2967
Qy 4093 t g c a g t g t c c c a t a a c c a t g a t g a t a a c a c a c t g c g g c c a a c t t a c t t c t g a c a a c g a t c 4152
Db 2968 T G C A G T G C T G C C A T A A C C A T G A T G A T A A C A C A C T C G G C C A A C T T A C T T C T G A C A A C G A T C 3027
Qy 4153 g g a g g a c g a a g g a g c t a a c c g c t t t t t t g c a a c a t a g g g g g a t c a t g t a a c t c g c c t 4212
Db 3028 G G A G G A C C G A A G G A G C T A A C C G C T T T T T T G C A C A A C A T G G G G G A T C A T G T A A C T C G C C T T 3087
Qy 4213 g a t c g t t g g g a a c c g g a g c t g a a t g a a g c c a t a c c a a a c g a c g a g c g t g a c a c c a c g a t g 4272
Db 3088 G A T C G T T G G G A A C C G G A G C T G A A T G A A G C C A T A C C A A A C G A C G A G C G T G A C A C C A C G A T G 3147
Qy 4273 c c t g t a g a a t g g c a a c a a c g t t g c g a a a c t a t t a a c t g g c g a a c t a c t a c t a g c t 4332
Db 3148 C C T G T A G C A A T G G C A A C A A C G T T G C G A A A C T A T T A A C T G G C G A C T A C T T A C T C T A G C T 3207
Qy 4333 t c c c g g c a a a t t a a t a g a c t g g a t g g a g g c g g g a t a a a g t t g c a g g a c c a c t t c t g c g c 4392
Db 3208 T C C C G G C A A C A A T T A A T A G A C T G G A T G A G G C G G G A T A A A G T T G C A G A C C A C T T C T G C G C 3267
Qy 4393 t c g g c c c t t c c g g c t g g c t g g t t a t t g c t g a t a a a t c t g g a c c g g t g a g c g t g g g t c t 4452
Db 3268 T C G G C C C T T C C G G C T G G C T G G T T A T T G C T G A T A A A T C T G G A G C C G G T G A G C G T G G G T C T 3327
Qy 4453 c g c g g t a t c a t t g c a g c a c t g g g g c c a g a t g g t a a g c c c t c c c g t a t c g t a g t a t c t a c 4512
Db 3328 C G C G G T A T C A T T G C A G C A C T G G G G C C A G A T G G T A A G C C C T C C C G T A T C G T A G T A T C T A C 3387
Qy 4513 a c g a c g g g a g t c a g g c a a c t a t g g a t g a a c g a a a t a g a c a g a t c g c t g a g a t g g t g c c 4572

Db 3388 A C G A C G G G G A G T C A G G C A A C T A T G G A T G A C G A A A T A G A C A G A T C G C T G A G A T A G G T G C C 3447
Qy 4573 t c a c t g a t a a g c a t t g g t a a c t g t c a g a c c a a g t t t a c t c a t a t a t a c t a c t t t a g a t t g a t 4632
Db 3448 T C A C T G A T T A A G C A T T G G T A A C T G T C A G A C C A A G T T T A C T C A T A T A T A C T T T A G A T T G A T 3507
Qy 4633 t t a a a a c t t c a t t t t t t a a t t t a a a a g g a t c t a g g t g a a g a t c c t t t t t g a t a a t c t c a t g 4692
Db 3508 T T A A A A C T T C A T T T T A A T T T A A A A G G A T C A G G T G A A G A T C C T T T T T G A T A A T C T C A T G 3567
Qy 4693 a c c a a a a t c c c t t a a c g t g a g t t t t c g t t c c a c t g a g c g t c a g a c c c g t a g a a a g a t c 4752
Db 3568 A C C A A A A T C C C T T A A C G T G A G T T T C G T T C C A C T G A G C G T C A G A C C C C G T A G A A A A G A T C 3627
Qy 4753 a a a g g a t c t t c t t g a g a t c c t t t t t t t t t c g c g c g t a a t c t g c t g c t t g c a a a c a a a a a a 4812
Db 3628 A A A G G A T C T T C T T G A G A T C C T T T T T T C T C G C G T A A T C T G C T G C T T G C A A A C A A A A A A 3687
Qy 4813 c c a c c g c t a c c a g c g g t g g t t t g t t t t t t t t t c g c g c g t a a t c t g c t g c t t g c a a a c a a a a a a g 4872
Db 3688 C C A C C G C T A C C A G C G G T G G T T T G T T T G C C G A T C A A G A G C T A C C A A C T C T T T T T C C G A A G 3747
Qy 4873 g t a a c t g g c t c a g c a g a g c g a g a t a c a a a a t a c t g t c c t t c t a g t a g c c g t a g t t a 4932
Db 3748 G T A A C T G G C T T C A G C A G A G C G C A G A T A C C A A A T A C T G T C C T T C T A G T G T A G C C G T A G T T A 3807
Qy 4933 g g c c a c c a c t t c a a g a a c t c t g t a g c a c c g c c t a c a t a c c t c g c t c t g c t a a t c c t g t t a 4992
Db 3808 G G C C A C C A C T T C A A G A A C T C T G T A G C A C C G C T A C A T A C C T C G C T C T G C T A A T C C T G T T A 3867
Qy 4993 c c a g t g g c t g c t c c a g t g g c g a t a a g t c g t c t t a c c g g t t g g a c t c a a g a c a g a t a g 5052
Db 3868 C C A G T G G C T G C T G C C A G T G G C G A T A A G T C G T G T C T T A C C G G G T T G G A C T C A A G A C G A T A G 3927
Qy 5053 t t a c c g g a t a a g g c g a g c g g t c g g g c t g a a c g g g g g t t c g t g c a c a c a g c c c a g c t t g 5112
Db 3928 T T A C C G G A T A A G C G C A G C G G T C G G G C T G A A C G G G G G T T C G T G C A C A C A G C C C A G C T T G 3987
Qy 5113 g a g c g a a c g a c c t a c a c c g a a c t g a g a t a c c t a c a g c g t g a g c t a t a g a g a a a g c g c c a c g 5172
Db 3988 G A G C G A A C G A C C T A C A C C G A A C T G A G A T A C C T A C A G C G T G A G C T A T G A G A A A G C G C C A C G 4047
Qy 5173 c t t c c c g a a g g g a g a a a g g c g g a c a g g t a t c c g g t a a g c g g a g g t c g g a a c a c a g g a g a g 5232
Db 4048 C T T C C C G A A G G G A G A A A G C G G A C A G A T A T C C G G T A A G C G G A C G G G T C G G A A C A C A G G A G A G 4107
Qy 5233 c g c a c g a g g a g c t t c c a g g g g g a a c c g c t g t a t c t t a t a g t c c t g t c g g g t t c g c 5292
Db 4108 C G C A C G A G G G A G C T T C C A G G G G A A A C C G C T G G T A T C T T T A T A G T C C T G T C G G G T T T C G C 4167
Qy 5293 c a c c t c t g a c t t g a g c g t c g a t t t t t g t g a t g c t c g t c a g g g g g c g a g c c t a t g g a a a 5352
Db 4168 C A C C T C T G A C T T G A C G C T G C A T T T T T G T G A T G C T C G T C A G G G G G C G G A G C C T A T G G A A A 4227
Qy 5353 a a c g c c a g c a a c g c g g c c t t t t a c g g t t c c t g g c c t t t t g c t g g c c t t t g c t c a c a t g 5412
Db 4228 A A C G C C A G C A A C G C G C C T T T T A C G G T T C C T G G C C T T T T G C T G G C C T T T T G C T C A C A T G 4287
Qy 5413 t t c t t t c c t g c g t t a t c c c c t g a t t c t g t g a t a a c c g t a t t a c c g c c t t t g a g t g a g c t 5472
Db 4288 T T C T T T C C T G C G T T A T C C C C T G A T T C T G T G G A T A A C C G T A T T A C C G C C T T T G A G T G A G C T 4347
Qy 5473 g a t a c c g c t c g c g a c c g a a c g a c c g a g c g a g c a g t c a g t a g a g a a g c g g a a 5532
Db 4348 G A T A C C G C T C G C G C A G C C G A A C G A C C G A G C G A G C G A G T C A G T G A G C G A G G A A G C G G A A 4407
Qy 5533 g a g c g c c a a t a c g a a a c c g c c t c t c c c c g c g c t t g c c g a t t c a t t a a t g c a g c t g g 5592
Db 4408 G A G C G C C A A T A C G C A A A C C G C C T C T C C C C G C G C T T G C C C G A T T C A T T A A T G C A G C T G G 4467
Qy 5593 c a c g a c a g g t t t c c c g a c t g g a a a g c g g g a g t g a g c g a a c g a a t t a a t g t a g t a g 5652
Db 4468 C A C G A C A G G T T T C C C G A C T G G A A A G C G G G C A G T G A G C G C A A C G C A A T T A A T G T G A G T T A G 4527

Qy 5653 ctcactcattagggcaccgccggttttacacatttatgcttcgggctcgctatgtgtgtgga 5712
|||||
Db 4528 CTCACTCATTAGGCCACCCAGGCTTTACACTTTATGCTTCGGGCTCGTATGTTGTGTGGA 4587

Qy 5713 attgtgagcggataacaatttcacacaggaacagctatgacctgattacgcca 5767
|||||
Db 4588 ATTGTAGCGGATAACAATTTACACACAGGAACAGCTATGACCATGATTACGCCA 4642

RESULT 14
AF276982
LOCUS AF276982 5027 bp DNA linear SYN 06-JUL-2000
DEFINITION Integrative vector pNLE1, complete sequence.
ACCESSION AF276982
VERSION AF276982.1 GI:8927566
KEYWORDS .
SOURCE Integrative vector pNLE1.
ORGANISM Integrative vector pNLE1
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 5027)
AUTHORS Salvatore,P., Cantalupo,G., Pagliarulo,C., Tredici,M., Lavitola,A.,
Bucci,C., Bruni,C.B. and Alifano,P.
TITLE A new vector for insertion of any DNA fragment into the chromosome
of transformable Neisseriae
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5027)
AUTHORS Salvatore,P., Cantalupo,G., Pagliarulo,C., Tredici,M., Lavitola,A.,
Bucci,C., Bruni,C.B. and Alifano,P.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2000) Biology, University of Lecce, Via
Monteroni, Lecce, LE 73100, Italy
FEATURES
source
1..5027
/organism="Integrative vector pNLE1"
/db_xref="taxon:130490"
/lab_host="Neisseria meningitidis"
/note="Integrative vector for transformable Neisseriae"
<32..733
/note="contains partial leucyl-tRNA synthetase (leuS)
coding region"
762..806
/note="neisserial uptake sequence and putative
transcription terminator"
970..>1513
/gene="dam"
970..>1513
/gene="dam"
/note="DNA methylase"
/codon_start=1
/transl_table=11
/product="DNA-adenine-methylase"
/protein_id="AAF82121.1"
/db_xref="GI:8927567"
/translation="MASKPFLKWAGGKHKLVPFIEHNLPTPTRKRLIEPFGSAALS
ALDFEYLLNDINADLIGLFRILKEKSGFIDYARSFISENNSDSRFYELREQNFS
KNLHERSALFIYLNHRHAFNGLCRYNSKGVNVPFGRYKSPYFPQEMEGFIQKSDRVE
LMCGDFQMILSQDNTDTVYC"
1605..2339
/gene="ermC"
1605..2339
/gene="ermC"
/note="RNA methylase"
/codon_start=1
/transl_table=11
/product="erythromycin resistance"
/protein_id="AAF82122.1"
/db_xref="GI:8927568"
/translation="MNEKNIKHSQNFITSKHNDIKIMTNIRLNEHDNIFEIGSGKH
FTELVQRCNFVTAIEDHKLCKTTENKLVHDNFDVNLKDILOFKFPKNQSYKIFGNI
PYNISTDIIRKIVFDSIADEIYLVIEYGFARLNTKRSALFLMAEVDISIISMVPR
EYFHPKPRVNSSLIRLNKKSRISHKDKQKYNFYVMKWVNKEYKKIFTKKNQFNLSLKH
AGIDDLNNISFEQFLSLFNSYKLFNK"

BASE COUNT 1424 a 1203 c 1127 g 1273 t
ORIGIN

Query Match 45.7%; Score 2634.4; DB 12; Length 5027;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3131 cgaattcactggccgctgcttttacaaacgctgctgactgggaaacccctggcgttacccaac 3190
|||||
Db 2392 CGAATTCACCTGGCCGTCGTTTTACAACTGCTGCTGACTGGGAAACCCCTGGCGTTACCCAAC 2451

Qy 3191 ttaatgccttgccagcacatcccccttccgagcctgaatggcgccctgatcgcggtatt 3250
|||||
Db 2452 TTAATCGCCTTGACGACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA 2511

Qy 3251 ccgatgccttccccaacagttgcgagcctgaatggcgccctgatcgcggtatt 3310
|||||
Db 2512 CCGATCGCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTGATCGGTATT 2571

Qy 3311 ttctccttacgcatctgtgcggtatttcacacccgcatatggtgcactctcagtacaatct 3370
|||||
Db 2572 TTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGTTGCTCATCTCAGTACAATCT 2631

Qy 3371 gctctgatgccgcatagtttaagccagcccgacacccgcgaaccccgctgacgcgcct 3430
|||||
Db 2632 GCTCTGATGCCGCATAGTTAAGCCAGCCCGACACCCGCCAACACCCGCTGACGGCCCT 2691

Qy 3431 gacggcttgcctgcctccggcatccgcttacagacaagctgtaccgtctccgggagct 3490
|||||
Db 2692 GACGGGCTGTCTGTCTCCGGCATCCGCTTACAGACAAGCTGTACCGCTCTCCGGGAGCT 2751

Qy 3491 gcatgtcagaggttttcacccgtcatcccgaaacgcgcgagacgaagggcctcgtga 3550
|||||
Db 2752 GCATGTGTCAGAGGTTTTTCACCGTCTACCGGAAACGCGGAGACGAAAGGGCCTCGTGA 2811

Qy 3551 tacgcctattttataggttaatgctcatgataataatggtttcttagacgtcaggtgcca 3610
|||||
Db 2812 TACGCCATTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCA 2871

Qy 3611 cttttcggggaaaatgtgcgcggaacccctatttgtttatttttctaataacattcaata 3670
|||||
Db 2872 CTTTTCGGGAAATGTGCGCGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATA 2931

Qy 3671 tgtatccgctcatgagacaataaacctcgataaaatgcttcaataatattgaaaaaggaaga 3730
|||||
Db 2932 TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAGGAAGA 2991

Qy 3731 gtagagtattcaacatttcctcgtgcgccttatcccttttttgcggcattttgccttc 3790
|||||
Db 2992 GTATGAGTATCAACATTTCCGTGTGCGCCTTATTCCCTTTTTTGGCGCATTTTGCCTTC 3051

Qy 3791 ctgttttgcctcaccagaaacgctggtgaaagtataaagatgctgaagatcagttgggtg 3850
|||||
Db 3052 CTGTTTTCCTCACCAGAAACGCTGCTGAAAAGTAAAAGATGCTGAAGATCAGTTGGGTG 3111

Qy 3851 cacgagtgggttacatcgaaactggatctcaacacgcggtaagatccttgagagttttcgcc 3910
|||||
Db 3112 CACGAGTGGGTTACATCGAACTGGATCTCAACACGCGGTAAGATCCTTGAGAGTTTTTCGCC 3171

Qy 3911 ccgaagaacgctttcccaatgatgagcacttttaaagttctgctatgtggcgcggtattat 3970
|||||
Db 3172 CCGAAGAACGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT 3231

Qy 3971 cccgtattgacgcggggaagagcaactcggctgcgcgcatacactattctcagaatgact 4030
|||||
Db 3232 CCCGTATTACGCCCGGGCAAGAGCAACTCGGTTCGCCGATACACTATTCTCAGAAATGACT 3291

Qy 4031 tggttgagtactcaccagtcacagaaaagcatcttacggatggcgatgacagtaagagaat 4090
|||||
Db 3292 TGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT 3351

Qy 4091 tatgcagtgctgccataaaccatgatgagtataaacactgcggccaacttacttctgacaacga 4150

Db 3352 TATGCAGTGCCTGCCATAAACCATAGTGATAAACACTGCGGCCAACTTACTTCTGACAAACGA 3411
QY 4151 tcggaggaccgaagagctaaaccgctttttgcaacaacatggggatcatgtaactcgcc 4210
Db 3412 TCGGAGGACCGAAGAGACTAACCGCTTTTTCACAAACATGGGGGATCATGTAACTCGCC 3471
QY 4211 ttgatcgttgggaaccggagctgaatgaagccataaccaaaacgacgagcgtgacaccacga 4270
Db 3472 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACGA 3531
QY 4271 tgcctgtagcaatggcaacaacgcttgccgaaaactattaactggcgaaactactactctag 4330
Db 3532 TGCCTGTAGCAATGGCAACAACGTTGCCGAAACTATTAACTGGCGAACTACTTACTCTAG 3591
QY 4331 cttcccggaacaataatagactggatggagcgggataaaagtbgcaggaccacttctgc 4390
Db 3592 CTTCCCGGCACAACATAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGC 3651
QY 4391 gctcgcccttcggtgctggtgttatgtctgataaaatctggagccggtgagcgtgggt 4450
Db 3652 GCTCGGCCCTTCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCGGTGAGCGTGGGT 3711
QY 4451 ctcgcggtatcatcgagcactgggcccagatggtaagccctcccgatcgtagtatct 4510
Db 3712 CTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT 3771
QY 4511 acacgacgggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtg 4570
Db 3772 ACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG 3831
QY 4571 cctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattg 4630
Db 3832 CCTCACTGATTAAGCATTGGTAACCTGCAGACCAAGTTTACTCATATATACTTTAGATTG 3891
QY 4631 atttaaaaacttcatttttaatttaaaaggatctagggtgaagatcccttttggataatctca 4690
Db 3892 ATTTAAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTGATAATCTCA 3951
QY 4691 tgacccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaaa 4750
Db 3952 TGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA 4011
QY 4751 tcaaaggatctcttgagatccctttttctgcgcgtaactcgtcgttgcaacaaaaa 4810
Db 4012 TCAAAGGATCTCTTGAGATCCCTTTTCTGCGGTAATCTGCTGCTGCAACAAAAA 4071
QY 4811 aaccaccgctaccagcggtgtgtttgttgcgggatcaagagctaccactcttttccga 4870
Db 4072 AACCACCGCTACCAGCGGTGTTGTTGCGGGATCAAGAGCTACCAACTCTTTTCCGA 4131
QY 4871 aggtaaactggcttcagcagagcgcagataccaaaatactgtccttctagtgtagccgtagt 4930
Db 4132 AGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTTCTTAGTGTAGCCGTAGT 4191
QY 4931 taggccaccacttcaagaactctgtagcaccgcctacatacctcgtctctgttaactcgtg 4990
Db 4192 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT 4251
QY 4991 taccagtggtcgtccagtgccgataaagtctgtcttacccgggttggaactcaagacgat 5050
Db 4252 TACCAGTGGCTGCTGCCAGTGGCGATAAAGTCGTGCTTTACCGGGTTGGACTCAAGACGAT 4311
QY 5051 agttaccggataaaggcgcagcggctcgggtgaacgggggttcgtgcacacagcccagct 5110
Db 4312 AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGTTTCGTGCACACAGCCCACT 4371
QY 5111 tggagcggaacgacctacacccgaactgagataacctacagcgtgagctatgagaaagcgcca 5170
Db 4372 TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 4431
QY 5171 cgcttcccgaaagggagaaaaggcggacaggtatcccggtgaagcggcaggggtcggaaacaggag 5230

Db 4432 CGCTTCCCCGAAGGGAGAAAGGGGACAGGTATCCGGTAAGCGGCACGGTTCGGAACAGGAG 4491
QY 5231 agcgcacagggagccttccaggggggaacgcctgggtatctttatagtcctgtcgggtttc 5290
Db 4492 AGCGACAGAGGAGCTTCCAGGGGGAACGCCTGGTATCTTTATAGTCCTGTTCGGGTTTC 4551
QY 5291 gccacctcgtacttgagcgtcgtgatttttctgtgatcctcgtcagggggcgagccctatgga 5350
Db 4552 GCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGAGCCTATGGA 4611
QY 5351 aaaaacgccaacgcgccctttttacggtttcctggcccttttctggccttttctgtcaca 5410
Db 4612 AAAACGCCAGCAACGCGCCTTTTACGGTTCTCGCTGGCCTTTTGTGGCCTTTTGCTCACA 4671
QY 5411 tgttcttctcgtgattatccccctgattctctgtggataaacctgattaccgcctttgagtga 5470
Db 4672 TGTTCCTTCTCGCTTATCCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG 4731
QY 5471 ctgataccgctcgcgcgacgcgaacgacgcgagcgcgagcgcgagtcagtcagcgcgagcgcg 5530
Db 4732 CTGATACCGCTCGCCGACGCCGAACGACGAGCGGACGAGTCAGTCAGGAGGAAAGCGG 4791
QY 5531 aagagcgcccaataacgcaaacgcctctccccgcgcgttggccgattcaatgaatgcagct 5590
Db 4792 AAGAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTCATTAAATGCAGCT 4851
QY 5591 ggcacgacaggtttcccgcactggaagcgagcgagcgcgagcgcgcaacgaattaatgtgagtt 5650
Db 4852 GGCACGACAGGTTCCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTT 4911
QY 5651 agctcactcattagggcaccccgctttcacactttatgcttccggctcgtatgtgtg 5710
Db 4912 AGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGSGCTCGTATGTTGTGTG 4971
QY 5711 gaattgtgagcggataacaatttcacacaggaagaaacagctatgacctgattacgcc 5766
Db 4972 GAATTGTGAGCGGATAACAATTTTCACACACAGAAACAGCTATGACCATGATTACGCC 5027

RESULT 15
BACKNB/c

LOCUS 5647 bp DNA linear BCT 14-APR-2000
DEFINITION Bacillus stearothermophilus pSTK1 and PUB110 genes for ORFs and kanamycin nucleotidyltransferase, complete cds.
ACCESSION D29979
VERSION D29979.1 GI:484046
KEYWORDS ORF; kanamycin nucleotidyltransferase; kanamycin resistance; thermostable mutation.
SOURCE Bacillus stearothermophilus DNA.
ORGANISM Geobacillus stearothermophilus
Bacteria; Firmicutes; Bacillus/Clostridium group; Geobacillus.
REFERENCE 1 (sites)
AUTHORS Liao,H., McKenzie,T. and Hageman,R.
TITLE Isolation of a thermostable enzyme variant by cloning and selection in a thermophile
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 83 (3), 576-580 (1986)
MEDLINE 86120983
PUBMED 3003740
REFERENCE 2 (sites)
AUTHORS Narumi,I., Nakayama,N., Nakamoto,S. and Kihara,H.
TITLE Complete nucleotide sequence of pSTK1, a cryptic plasmid from Bacillus stearothermophilus TK015
JOURNAL Biotechnol. Lett. 15, 1013-1016 (1993)
REFERENCE 3 (sites)
AUTHORS Narumi,I., Nakayama,N., Nakamoto,S., Kimura,T., Yanagkisawa,T. and Kihara,H.
TITLE Construction of a new shuttle vector pSTE33 and its stabilities in Bacillus stearothermophilus, Bacillus subtilis, and Escherichia coli
JOURNAL Biotechnol. Lett. 15, 815-820 (1993)
REFERENCE 4 (sites)

AUTHORS Narumi,I., Nakayama,N., Nakamoto,S. and Kihara,H.
TITLE Bacillus stearothermophilus plasmid pSTK1 replicon is functional in Escherichia coli
JOURNAL Biotechnol. Lett. 17, 475-480 (1995)
REFERENCE 5 (bases 1 to 5647)
AUTHORS Nakamoto,S.
JOURNAL Unpublished (1995)
FEATURES Location/Qualifiers
source 1. .5647
/organism="Geobacillus stearothermophilus"
/db_xref="taxon:1422"
CDS 243. .518
/note="pSTK1"
/citation-[2]
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF1"
/protein_id="BAA06246.1"
/db_xref="GI:849038"
/translation="MKKIQIRLTEENDKKLREIAERYGMSINSVASFIIGKWLDENYD LKDRILEGLAASQNEVMERVAQVVEELFSPNEKIKELERHGLKLN"
rep_origin 524. .713
/note="pala minus"
/citation-[2]
/evidence=not_experimental
CDS 538. .768
/note="pSTK1"
/citation-[2]
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF2"
/protein_id="BAA06247.1"
/db_xref="GI:849039"
/translation="MPRKLLARVRRVLGWTSRKALPCICVQAFSRQRLPAPADKHASR WGPVPKGNWNNAERPVIGGIKIGNSESTPVV"
CDS 1013. .1822
/note="pSTK1"
/citation-[2]
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="ORF3"
/protein_id="BAA06248.1"
/db_xref="GI:849040"
/translation="MGIHVEMTGQGCRLFELHTSINWYELFYRLVYEVNITRLDVA VDDFKGYFKINTLVKKLKDDVTSRFKKARHIENIVIEGETIGHTLYFGAPSSDIQV RFYEKNVQMGMDIDVWNRTIQLRDDRAHVVAQIIADDVLPGLGEIVAGLLRNYIQFRT RKATDNKKRWPLARFWLNFGLDVQPLRIAKQMPKTSIEKKYRWIDSQVSKSFFMIYY CLNEEEKQRFIDDLAEGASKLTKADLQVINQFQFSKNITYDEMIKIIROSK"
CDS 2168. .2929
/note="kanamycin resistance gene
pUB110"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="kanamycin nucleotidyltransferase"
/protein_id="BAA06249.1"
/db_xref="GI:484047"
/translation="MNGPIIMTREERMKIVHEIKERILDKYGGDDVKAIGVYGSGLGRQT DGPYSDIEMCMVNSTEEAEFSHEWTTGEWKVEVNFYSEEILLDYASQVESDWPLTHGQ FFSILPIYDSGGYLEKVYQTAKSVEAQKFHDAICALIVEELFEYAGKWRNIRVQGPPT FLPSLTVQVAMAGAMLIGLHHRICYTTSASVLTEAVKQSDLPSPGYDHLCCQFVMSGQLS DSEKLLSLENFWNGIQEWTERHGYIVDVSKRIPP"
variation 2405
/citation-[1]
/phenotype="thermostable mutant"
/replace="g"
variation 2556
/citation-[1]
/phenotype="thermostable mutant"
/replace="c"

BASE COUNT	1579 a	1149 c	1470 g	1449 t	
ORIGIN					
Query Match	45.7%	Score 2633.8;	DB 1;	Length 5647;	
Best Local Similarity	99.9%	Pred. No. 0;			
Matches 2635;	Conservative	0;	Mismatches	2;	Indels 0; Gaps 0;
Qy	3131	cgaattcaactggccgctgctttttacaacagtcgctgactgggaaaaacccctggcgttaccccaac	3190		
Db	5635	CGAATTCACTGGCCGTCGTTTACACGTCGTGACTGGGAAAACCCCTGCGTTACCCAAC	5576		
Qy	3191	ttaatgccttgacgacacatcccccttccgacgctggcgtaataagagaagagcccgca	3250		
Db	5575	TTAATCGCCTTGACGACACATCCCCCTTTGCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCA	5516		
Qy	3251	ccgacgccttcccaacagttgcgcagcctgaatggcgatggcgctgatcggtatt	3310		
Db	5515	CCGATCGCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGCGCTGATCGGTATT	5456		
Qy	3311	ttctccttaagcatctgtgcggtattttcacacccgcataatggtgcactctcagtacaatct	3370		
Db	5455	TTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATATGGTGCACCTCTCAGTACAATCT	5396		
Qy	3371	gctctgatgcgcgcatagttaagccagccccgcgacacccgcgaacacccgcgtgacgcgccct	3430		
Db	5395	GCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCT	5336		
Qy	3431	gacgggtctgtctgtctcccggcacccgcttacagacaagctgtgacgctctccgggagct	3490		
Db	5335	GACGGGCTGTCTGTCTCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCT	5276		
Qy	3491	gcattgtcagaggttttcaccgctcatcaccgaaacgcgcgagacgaaaggcctcgtga	3550		
Db	5275	GCATGTGCAGAGGTTTTCACCGTCTACCGGAAACGCGGAGACGAAAGGGCCTCGTGA	5216		
Qy	3551	tacgcctatttttataggttaattgatcatgataataatggtttcttagacgctcaggtgcca	3610		
Db	5215	TACGCCATTTTATAGGTTAATGTCTATGATAATAATGGTTTCTTAGACGTCAGGTGGCA	5156		
Qy	3611	cttttcggggaaaatgtgcgcgggaacccctatttggttattttcttaaatcacattcaataa	3670		
Db	5155	CTTTTCGGGAAATGTGCGGGAACCCCTATTGTTTATTTTCTAAATACATTCAAATA	5096		
Qy	3671	tgtatccgctcatgagacaataaacccctgataaattgcttcaataataattgaaaaaggaaga	3730		
Db	5095	TGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATAATTGAAAAAGGAAGA	5036		
Qy	3731	gtatgatttcaacattcccgctgtgcgccttattcccttttttggcgcatatttgccttc	3790		
Db	5035	GTATGATATTCAACATTTCCGCTGTGCGCCTTATTCCCTTTTGGGCGCATTTTGCCTTC	4976		
Qy	3791	ctgttttgcctcaccagaaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg	3850		
Db	4975	CTGTTTGTCTCACCAGAAACGCTGCTGTAAGTAAAGATGCTGAAGATCAGTTGGGTG	4916		
Qy	3851	cacgagtggttacatcgaaactggatctcaacacgcggtaagatcccttgagagttttcgcc	3910		
Db	4915	CACGAGTGGGTTACATCGAACTGGATCTCAACACGCGGTAAGATCCTTGAGAGTTTTCGCC	4856		
Qy	3911	ccgaagaacgttttccaatgatgagcacttttaagttctgctatgtgtgcgcggtattat	3970		
Db	4855	CCGAAGAAGCTTTTCCAATGATGAGCATTTTAAAGTTCTGCTATGTGGCGCGGTATTAT	4796		
Qy	3971	cccgattgacgcggggaagagcaactcggtcgccgcatacactatttctcagaatgact	4030		
Db	4795	CCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAATGACT	4736		
Qy	4031	tggttgagtactcaccagtcacagaaaaagcatcttacggatggcgatgacagtaagagaat	4090		
Db	4735	TGTTTGAGTACTCACCAGTCACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAAT	4676		
Qy	4091	tatgcagtgctgccataaaccatgatgtgataaacactgcggccaacttacttctgacaacaga	4150		

Db	4675	 TATGCAGTGCTGCCATAACCATAGAGTGATAACACTGGCGCCAACTTACTTCTGACAACGA	4611
QY	4151	tcggaggaccgaaggagcgtaacccgcttttttgcacaacatgggggatcatgtaaactcgcc	4210
Db	4615	TCGGAGGACCGAAGGAGCTAACCGCTTTTGTGCAACAACATGGGGGATCATGTAACTCGCC	4556
QY	4211	ttgatcgttgggaaccgagcgtgaatgaagccataaccacaaacgacgagcgtgacaccacga	4270
Db	4555	TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA	4496
QY	4271	tgccctgtagcaatggcaacaacgcttgcgcgaaactattaaactggcgaaactacttactctag	4330
Db	4495	TGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAG	4436
QY	4331	cttccccggcaacaataatagactggatggaggcggaataaaagtgcaggacacattctgc	4390
Db	4435	CTTCCCGGCAACAATTAATAGACTGGATGGAGGGGATAAAAGTTGCAGGACCACTTCTTGC	4376
QY	4391	gctcgcccttccggctcggctggtttatttgcgtgataaaactctggagccggtgagcgtgggt	4450
Db	4375	GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGTATAAATCTGGAGCCGGTGAGCGTGGGT	4316
QY	4451	ctcgcggtatcattgcagcactggggccagatggtaagccctcccgatcgttagtttatct	4510
Db	4315	CTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT	4256
QY	4511	acacgacggggagtcaggcaactatggatgaacgaaaaatagacagatcgcctgagataggtg	4570
Db	4255	ACACGACGGGGAGTCAGCAAACTATGGATGAACGAAAATAGACAGATCGCTGAGATAGGTG	4196
QY	4571	cctcactgattaaagcattggttaactgtcagaccgaagtctactcatatatacttttagattg	4630
Db	4195	CCTCACTGATTAAAGCATTTGGTAACCTGTGAGACCAAGTTTACTCATATATATCTTAGATTG	4136
QY	4631	atttaaaaacttcatttttaatttaaaggatctagggtgaagatccttttttgataaatctca	4690
Db	4135	ATTTAAACTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCCTTTTGTGATAATCTCA	4076
QY	4691	tgacaaaaatcccttaacgctgagtttcttcactgagcgtcagaccocgtagaaaaaga	4750
Db	4075	TGACCAAAATCCCTTTAACGTGAGTTTTCGTTCCACTGAGCGTTCAGACCCCGTAGAAAAAGA	4016
QY	4751	tcaaaggatcttcttgagatccctttttctggcgtaaatctgctgcttgcaaaacaaaaa	4810
Db	4015	TCAAAGGATCTTCTTGAGATCCCTTTTCTCGCGGTAATCTGCTGCTTGCAAAACAAAAA	3956
QY	4811	aaccaccgctaccagcggtgttgtttgcccggatcaagagctaccaactctttttccga	4870
Db	3955	AACCACCGCTACCGCGGTGGTTGTTTTGCCGGATCAAGAGCTACCAACTCTTTTCCGA	3896
QY	4871	aggtaaactggcttcagcagagcgcagataacaaaatactgtccttctagtgtgtagccgtagt	4930
Db	3895	AGGTAACTGGCTTCAGCAGAGCGCAGATACCAAAATACTGTCTCTTCTAGTGTAGCCGTAGT	3836
QY	4931	taggccaccacttcaagaactctgtagcacgcctacataccctcgctctgctaaatcctgt	4990
Db	3835	TAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCCTGT	3776
QY	4991	taccagtggctgctgccagtggcgataaagtctgtcttaccggggttgagactcaagacgat	5050
Db	3775	TACCAGTGGCTGCTCCAGTGGCGATAAGTCTGTCTTACC GG GTTGGACTCAAGACGAT	3716
QY	5051	agttaccgggataaaggcgagcggtcggttgaaacgggggttcgtgcacacagccccagct	5110
Db	3715	AGTTACCGGATAAGGCGCAGCGGTTCGGCTGAACGGGGGGTTCGTGACACAGCCCCAGCT	3656
QY	5111	tggagcgaacgacctacacccaactgagataacctacagcgtgagctatgagaaagcgcca	5170
Db	3655	TGGAGCGAACGACCTACACCGAACTGAGATAACCTACAGCGTGAGCATTGAGAAAGCGCCA	3596
QY	5171	cgcttccccgaaggagaaaggcgacaggtatccggtaagcggcgaggttcggaacaggag	5230

•

Search completed: August 31, 2002, 18:10:48
Job time: 16101 sec